Efforts Towards the Total Synthesis of Labelled CCL2 Proteins and Dipeptide Chemotaxis Inhibitors

HUDSON, ALEXANDER,STEPHEN

Efforts Towards the Total Synthesis of Labelled CCL2 Proteins and Dipeptide Chemotaxis Inhibitors

A Thesis submitted for the degree of

Doctor of Philosophy

By

Alexander Stephen Hudson

&

Department of Chemistry

Durham University

November 2015

Abbreviations

7TM – Seven Transmembrane
ABI – Applied Biosystems
Acm – Acetamidomethyl
Alloc – Allyloxycarbonyl
ArX – Halo-aromatic
BBB – Blood Brain Barrier
BFTA – 3-bromo-1,1,1-trifluoroacetone
BMC – Bare Membrane Chemotaxis
Boc – *t*-Butyloxycarbonyl
BSA – Bovine Serum Albumin
Bzl – Benzyl
CAM – Cell Adhesion Molecule
CCL2 – CC Motif Chemokine Ligand 2
COPD – Chronic Obstructive Pulmonary Disease
dba – Dibenzylideneacetone
Dbz – 3,4-Diaminobenzoic acid
DCC – *N*,*N*-dicyclohexylcarbodiimide
DCM – Dichloromethane
DIC –*N*,*N*-Diisopropylcarbodiimide
DIPEA – *N*,*N*-Diisopropylethylamine
DKP – 2,5-Diketopiperazine
DMAP – 4-(dimethylamino)pyridine
DMF – *N*,*N*-Dimethylformamide
DMSO – Dimethyl Sulfoxide
EDG – Electron Donating Group
EDT – 1,2-Ethanedithiol
ESI – Electrospray Ionisation
EWG – Electron Withdrawing Group
FDA – Food and Drug Administration
Fmoc – Fluorenylmethyloxycarbonyl
GAG – Glycosaminoglycan
Gdn – Guanidine
GPCR – G Protein Coupled Receptor
GTP – Guanosine-5'-triphosphate
GVHD – Graft Versus Host Disease
HBTU – *O*-(Benzotriazol-1-yl) *N*,*N*,*N*,*N*-tetramethyluronium hexafluorophosphate
Hept – Heptyl
HIV-1 –Human Immunodeficiency Virus-1
HL – High loading
HMEC-1 – A Human Microvascular Endothelial Cell Line
HOBt – Hydroxybenzotriazole
IR – Infrared
LCMS – Liquid Chromatography Mass Spectrometry
LG – Leaving Group
LL – Low Loading
RNS – Reactive Nitrogen Species
ROS – Reactive Oxygen Species

RP-HPLC – Reverse Phase High Performance Liquid Chromatography
mAb – Monoclonal Antibody
MALDI-TOF – Matrix Assisted Laser Desorption/Ionization Time of Flight
MeCN – Acetonitrile
MeOH – Methanol
MESNa – 2-Mercaptoethanesulfonate Sodium Salt
MHBA – 4-Methylbenzhydrylamine
MMP – Matrix Metalloproteinase
MPAA – (4-Carboxymethyl) Thiophenol
MS – Mass Spectrometry
MSMS – Tandem Mass Spectrometry
NADPH – Nicotinamide Adenine Dinucleotide Phosphate
Nbz – *N*-Acylbenzimidazolinone
NCL – Native Chemical Ligation
NMP – *N*-Methyl-2-pyrrolidinone
NMR – Nuclear Magnetic Resonance
PAM – 4-Hydroxymethyl-phenylacetamidomethyl Resin
Pbf – 2,2,4,6,7-Pentamethyl-dihydrobenzofuran-5-sulfonyl
PBS – Phosphate Buffered Saline
PBSF – Perfluoro-1-butanesulfonyl Fluoride
PCM – Post Condensational Modification
PEG – Poly(ethylene glycol)
PS – Polystyrene
PyBOP – (Benzotriazol-1-yloxy)tripyrrolidinophosphonium Hexafluorophosphate
PyIRS – Pyrrolysyl-tRNA Synthetase
RCM – Ring Closing Metathesis
RPMI – Roswell Park Memorial Institute
SEA – Bis(2-sulfanylethyl)amino
SPPS –Solid Phase Peptide Synthesis
TBAT – *t*-Butyl Ammonium Difluorophenyl Silicate
Tbfmoc –Tetrabenzo[a,c,g,i] Fluorenyl-17 Methoxycarbonyl
TBTA – *t*-Butyl 2,2,2-trichloroacetimidate
*t*Bu – *t*-Butyl
TCEP – Tris(2-carboxyethyl)phosphine
TEC – Transendothelial Chemotaxis
TFA –Trifluoroacetic Acid
TFET – 2,2,2-Trifluoroethanethiol
THF – Tetrahydrofuran
THP-1 – A Human Acute Monocytic Leukemia Cell Line
Thz – 1,3-Thiazolidine4carboxyl
TIPS – Triisopropylsilane
TMS – Trimethylsilyl
TNBS – 2,4,6-Trinitrobenzenesulfonic
tris –Tris(hydroxymethyl)aminomethane
tRNA – Transfer Ribonucleic Acid
Trt – Trityl
TyrRS - Tyrosyl-tRNA Synthetase
UV – Ultraviolet
Wang/PABA – *p*-Amino Benzoic Acid
Xaa – Any Amino Acid

Abstract

Chemokines such as CCL2 are small proteins with molecular weights between 8-10 kDa. They promote chemotaxis and play a vital role in the recruitment of leukocytes to the site of inflammation. Given their key biological functions, understanding their mechanism of action and inhibiting their action has therapeutic potential in a range of diseases.

Selective inhibitors of CCL2 induced chemotaxis based on the diketopiperazine (DKP) natural product, cyclo(13,15-dichloro-L-Pro-L-Tyr) were recently reported by the Cobb group. In order to develop this work further and to produce an expanded library, we optimised an on-resin DKP synthesis. In collaboration with researchers at Newcastle Medical School, chemotaxis assays were performed in an attempted to define the structural features (required for inhibition) of the DKPs. To faciliate the aforementioned work, facile synthetic routes to a number of novel heteraromatic and proline derivatives were developed.

The posttranslational nitration of CCL2 is used as a mechanism to a modulate CCL2. However, in all previous studies nitrated CCL2 was utilised as a heterogeneous mix of protein. Having access to single site nitrated CCL2 will enable the mechanism of abrogration to be pinpointed to a specific residue. Therefore, the total chemical synthesis of both native CCL2 and site-specifically nitrated CCL2 *via* solid phase peptide synthesis (SPPS) was undertaken. The work focused on the potential application of microwave assisted SPPS as a means to rapidly access the target proteins.

Memorandum

This thesis describes work carried out within the Chemistry Department at Durham University and at Cambridge Research Biochemicals Ltd (Billingham, UK) between July 2011 and June 2015. Unless acnkowleged by reference, the work described is entirely of the author and has not been submitted for any other degree. Quotation from this thesis should only be published with the authors written consent and copyright rests with the author. Information derived from this thesis should be properly acknowledged.

The work contained in this thesis has been presented, in part, at;

RSC Fluorine Chemistry Conference (Aberdeen). (8 Sept 2011).

Protein and Peptide Science Group Early Stage Researcher Meeting (RSC London). (8 Nov 2011)

3rd RSC Bio-Organic Group Forum (Leicester). (5 July 2012).

Protein and Peptide Science Group Early Stage Researcher Meeting (RSC London) (29 Nov 2012).

Durham University Postgraduate Symposium (Durham). (2 April 2013).

23rd American Peptide Symposium (Hawaii). (22 June 2013).

Protein and Peptide Science Group Early Stage Researcher Meeting (Durham). (Nov 2013).

Durham University Postgraduate Symposium (Durham). (18 June 2014)

Probing Molecular Interactions and Mechanism at the Chemistry/Biology Interface (Durham). (7 July 2014)

Chemistry and Biology of Peptides (NIMR London). (18 July 2014)

The work contained in this thesis has been published, in part, in the following journal articles;

A.S. Hudson, L Caron, N Colgin, and S.L Cobb. A direct method for the synthesis of orthogonally protected furyl- and thienyl-amino acids. *Amino Acids*, **2015**, 7(4):779-85.

N.K. O'Connor, A.S. Hudson, S.L. Cobb, D O'Neil, J Robertson, V Duncan and C.D. Novel fluorinated lipopeptides from *Bacillus* sp. CS93 *via* precursor-directed biosynthesis. *Amino Acids*, **2014**, 46(12):2745-52.

Acknowledgements

I would like to thank my supervisor Dr Steven Cobb at Durham University for being the major catalyst for me to undertake postgraduate research. His advice and encouragement throughout this PhD have not only made the process possible but enjoyable. From a small group in 2010 I have witnessed the effort and compassion required to make a research group not only successful but grow and flourish. Being part of the Cobb group has been a pleasure and I have countless happy memories from every past and present member; Dr Asahi Cano-Marques for football and fun, Dr Francis Chadbourne for easing my first steps into a research lab, Sam Lear, Lara Small, Dr Bnar Jawdat, Maria Czyzweska, Caitlin Mooney, Hannah Bolt, Alex Webster, Dr Chris Coxon, Dr Anika Dose, Dianna Giménez**,** Nicholas Robson, Ambrose Crofton, Malgorzata Przeradzka, Andrew Steer, Mark Laws, Ed Rendle, Miriam Everitt, Roderick Dirkzwager, and Edward Jones. There are a number of other colleagues that deserve thanks as many of them have improved my life in many ways. Rose Simnett for the enduring support and joy she has given me, Dr Davey Cole, Dr Paul Brooks, Dr Serena Agostini, Brunella Maranesi, Tatiana Lovato, Dr Peter King, Cat Blackwell, Sercan Sevinc, Kieran Atter, Oliver Burnham, Russel Balster, Andrew Longstaff, Dr Iain Johnson, Shengui Hou, Dr Scott Sadler, Oliver Maguire, Dr Adam Hope, Dr Ffion Abraham, Dr Rob Goodwill, Vicki Linthwaite and Dr Rebek Pal.

I would also like to thank my co-supervisor Dr Laurent Caron at Cambridge Research Biochemicals. His knowledge of peptide chemistry and kind guidance has helped me throughout. I am indebted to Emily Humphries and Alison White who as directors at Cambridge Research Biochemicals have provided funding as well as the opportunity to work in a wonderful company full of brilliant staff. Thank you; Jodie Fleming, Dr Sophie Clift, Dr Natasha Lethbridge, Jess Tomlinson, Toby Stephenson, Elie Coutts, Hannah Jones, Elizabeth Law, Sarah Wallace, Claire Burn, Jessica Kane-Fidgeon, David Brooks. I would like to give special thanks to Dr Neil Colgin, who I have had the privilege to work with during both his post-doctoral position in the Cobb group and as chemist at Cambridge Research Biochemicals. Neil has shared his passions for organic chemistry, great music and has consistently gone above and beyond to support my research.

Dr Simi Ali at Newcastle University and Dr Cormac Murphy at University College Dublin and their respective research groups deserve my thanks. Notably, Dr Catriona Barker and Dr Neil O'Connor: Catriona for her patience and guidance in biology & folk music and Neil for the scientific collaboration as well as sharing many similar interests and becoming a great friend.

My family have been exceptionally supportive and given me opportunity to achieve throughout my life and education. They've sparked my academic endeavours, supported my sporting aspirations (with the countless trips to football games in the dreariest of weather) and provided a calm and loving base. To mention but a few: My parents, Julie and Stephen for their unconditional patience, encouragement, happiness and love and my sister Dominique for her saintly tolerance. My grandparents: Kath and Lawrence for their love and constant supply of braised steak and stories. My cousins: John and Callum, for always being close friends and my aunts, Georgina and Johann. Thank you all.

Finally, during my time as a postgraduate I have met some exceptional people in the sporting aspect of Durham University and I feel these lasting friendships deserve my acknowledgements: Matthew Bainbridge, Dr Lee Weller and Dr Renan Petersen-Wagner for sharing so many of their worldly experiences and helping create many memories together. The three of them, amongst others have helped shape Ustinov AFC to be a constant source of happiness and a pleasure to myself and hundreds of other postgraduates. I feel privileged to be a part of this club and thus, I would like to thank some of the friends I have made that have not previously been acknowledged: Phillip Gill, Alex Martin, Alex Nicholls, David Williams, Jon Worrall, Dr Andy Bonfini, Eddy Walter, Chris Dobson, Daniel Conway, Dr Marc Owen-Jones, David Hyde, Daniel Holck and Bruce Rawlings.

Durham University

crb CAMBRIDGE RESEARCH BIOCHEMICALS

RESEARCH GROUP

Contents

Chapter 1: Introduction

1.1 A General Introduction to Proteins p2

1.2 Chemokines p4

1.2.1 Introduction to Chemokines p4

1.2.2 Classification of Chemokines p5

1.2.3 The Simplified Mode of Action of Inflammatory Chemokines p7

1.2.4 The Specific and Non-Specific Nature of Chemokines p8

1.3 Chemokine Mediated Immune Response p14

1.3.1 Inflammation p14

1.3.2 Chronic Inflammation p15

1.3.3 Oxidative Stress p15

1.3.4 Peroxynitrate p17

1.3.5 Nitration of Chemokine: CCL2 p19

1.4 Chemokines as Biological Targets p21

1.4.1 Therapeutics to Target Chemokine/Chemokine Receptor Interactions p21

1.4.2 Small Molecule Chemokine Inhibitors: Acylamino-Lactams p23

1.4.3 Diketopiperazine Chemokine Inhibitors from Natural Products p24

1.4.4 Results of the In Vitro Studies of the Chemokine Inhibitors p25

1.5 Aims p27

1.5.1 Site-Selective Studies of the Post Translational Nitration of CCL2 p27

1.5.2 Increased Studies on DKP Based CCL2 Specific Inhibitors p29

1.6 References p30

Chapter 2 The Synthesis and Applications of Novel Aromatic & Heteroaromatic Amino Acids

2.1 Introduction p35

2.1.1 Expanding Nature's "Tool-box" of Amino Acids p35

2.1.2 Incorporation of Novel/Unnatural Amino Acids into Peptides and Proteins p37

2.1.3 Previous Syntheses of Unnatural α-Amino Acids p39

2.1.4 Unnatural Amino Acid Synthesis via Pd-catalysed Negishi Cross-Coupling Synthesis p41

2.2 Bioactive 5-Membered Heteroaromatic Molecules p44

2.2.1 Furyl and Thienyl Moieties in Drug Molecules and Natural Products p44

2.3 Synthesis of 5-Membered Heteroaromatic Amino Acids via Pd-Catalysed Negishi Cross-Coupling. p45

2.3.1 Previous Synthesis of 5-Membered Heteroaromatic Amino Acids p45
2.3.2 Preparation of Starting Materials p47
2.3.2.1 Halo-Aromatic p47
2.3.2.2 β-Iodo Amino Acid p48
2.3.3 Syntheses of Heteroaromatic N-Fmoc Amino Acids via Pd-Catalysed Negishi Cross-Coupling p51
2.3.4 Optimisation of the Synthesis of Heteroaromatic N-Fmoc Amino Acids p53
2.3.5 Synthesis of Heteroaromatic N-Boc Protected 2-furyl Alanine p60
2.3.6 Synthesis of γ-Glutamyl-2-furylalanine p60
2.3.7 Heteroaromatic Amino Acids Conclusions p63
2.4 Biologically Relevant Fluorine Containing Molecules p64
2.4.1 Organofluorine in Medicinal Chemistry p64
2.4.2 Fluorine in Peptide/Protein Chemistry p65
2.4.3 Iturin A and Lipopeptides from Bacillus sp. CS93 p68
2.5 Synthesis of Fluoro-tyrosine Derivative Amino Acids via Pd-Catalysed Negishi Cross-Coupling. P69
2.5.1 Previous Syntheses of Fluorinated Tyrosine Derivative Amino Acids p69
2.5.2 Preparation/Choice of Starting Materials p70
2.5.2.1 Synthesis of β-Iodo Amino Acid p70
2.5.2.2 Preparation of Halo-Aromatic p70
2.5.3 Synthesis of Fluoro-Tyrosines p71
2.5.4 Biosynthesis of Fluoro-Tyrosine Lipopeptides: Iturin A and Fengycin p72
2.5.5 Fluorinated Tyrosine Conclusions p74
2.6 General Conclusions on Pd-Catalysed Cross-Coupling Heteroaromatic Synthesis p74
2.7 References p75

Chapter 3: The Synthesis and Applications of Thiol Containing Amino Acids in NCL
3.1 β-Thiol Containing Amino Acids p80
3.1.1 Native Chemical Ligation (NCL) p80
3.1.2 Native Chemical Ligation Limitations p82
3.1.3 Multiple Ligation Strategies: N-Terminal Cysteine Protection p83
3.1.4 Multiple Ligation Strategies: Kinetic Native Chemical Ligation p84
3.2 Syntheses of Thiol Containing Amino Acids p86
3.2.1 Synthesis of a Small Library of Thiol Containing Amino Acids p86
3.2.1.1 Thioproline Syntheses: Cyclic Sulfamidate p88
3.2.1.2 Thioproline Syntheses*:* Tetrafluoropyridine Leaving Group p89

3.2.1.3 Thioproline Syntheses: Mesylate Leaving Group p90
3.2.1.4 Thioproline Syntheses: Iodine Leaving Group p91
3.2.2 Thiol Amino Acid Synthesis Conclusions p93
3.3 Applications of Thiol amino acids p94
3.3.1 pK_a Studies p94
3.3.2 pK_a Studies: Conclusions p97
3.4 References p98

Chapter 4: Diketopiperazines - Synthesis and Inhibitory Activity Against CCL2 Induced Chemotaxis.
4.1 An Introduction to 2,5-Diketopiperazines p100
4.1.1 2,5-Diketopiperazines: A Structural Overview p100
4.1.2 DKPs as Inhibitors of Chemokine Mediated Chemotaxis p103
4.2 Introduction to the Synthesis of DKPs p107
4.2.1 Solution Phase Dipeptide Ester Cyclisation Synthesis of DKPs p108
4.2.3 On-Resin Cyclisation and Release Synthesis of DKPs p110
4.3 Preparation of a New Library of DKPs p111
4.4 On-Resin Synthesis of a 2nd Generation of DKPs p114
4.4.1 Resin Loading and Synthesis of the Linker p114
4.4.2 On-Resin Synthesis of the DKP Library p115
4.5 Biological Studies of the New Library DKP p117
4.5.1 CCL2 Induced Chemotaxis: In Vitro p117
4.5.2 Chemotaxis Library Screen via a BMC Assay p119
4.5.3 Anti-Fungal Activity p122
4.6 Conclusions and Analysis of the DKP Inhibitor Library p122
4.6.1 Stereochemistry p122
4.6.2 Aromatic Moiety p123
4.6.3 Proline Moiety p124
4.7 Library Expansion Using Unnatural Aromatic and Heteroaromatic Amino Acids p124
4.7.1 Target Molecules p124
4.7.2 Syntheses of Cyclo(flp-L-Phe) and Cyclo(L-Pro-2-furyl-L-Ala) p126
4.8 General Conclusions p128
4.9 References p129

Chapter 5: Attempted Syntheses of the Native Human Chemokine CCL2 and its Analogues
5.1 Protein Synthesis p132
5.1.1 Chemical Synthesis of Peptides/Proteins p132
5.1.2 SPPS of Peptides/Proteins p132
5.1.3 Native Chemical Ligation (NCL) p134
5.1.4 Thioester and Solid Support p134
5.2 Previous Syntheses of Human Chemokine CCL2 p137
5.2.1 The First Total Chemical Synthesis of CCL2 p138
5.2.2 A Linear Stepwise N-Fmoc Total Chemical Syntheses of CCL2 p138
5.2.3 Total Chemical Synthesis of CCL2 via NCL p140
5.2.4 Preliminary Work on the Microwave Assisted Linear Synthesis of CCL2 p143
5.3 Microwave Assisted Total Chemical Synthesis of CCL2: A Fragmentation Approach p145
5.3.1 A Fragment Approach to the Total Synthesis of CCL2 p145
5.3.1.1 Synthesis of CCL2 Peptide Fragment 52-76 p148
5.3.1.2 Synthesis of Peptide Fragment 36-51 p149
5.3.1.3 NCL of CCL2 Peptide Fragments 36-51 and 52-76 p152
5.4 Microwave Assisted Total Chemical Synthesis of CCL2: An Optimised Stepwise Approach p154
5.4.1 An Optimised Stepwise Approach to the Total Synthesis of CCL2 p154
5.4.2 Synthesis of CCL2 Peptide Fragment 1-35 p160
5.4.3 NCL of CCL2 Peptide Fragment 1-35 and 36-76 p168
5.5 Synthesis of Nitrated Tyrosine CCL2 p169
5.5.1 Incorporation of a Nitrated Tyrosine in CCL2 p169
5.6 Conclusions p173
5.7 References p174

Chapter 6: Conclusions and Future Work
6.1 Conclusions and Future Work p177
6.1.1 Amino Acid Syntheses p177
6.1.2 DKPs and CCL2 Induced Chemotaxis Inhibition p180
6.1.3 Potential Routes to the Microwave Assisted Synthesis of CCL2 p184

Chapter 7: Experimental
7.1 Small Molecule Syntheses p187
7.1.1 Materials and Methods p187

7.1.2 General Procedures p187
7.1.2.1 Synthesis of Orthogonally Protected Heteroaromatic Amino Acids via Pd-Catalysed Negishi Cross-Coupling p187
7.1.2.2 Solid Phase Cyclic Dipeptide (DKP) Synthesis p188
7.1.3 Small Molecule Syntheses p189
7.2 Solid Phase Peptide Syntheses p213
7.2.1 Materials and Methods p213
7.2.2 General SPPS Procedures p214
7.2.2.1 Microwave Assisted Automated Fmoc-SPPS p214
7.2.2.2 Room Temperature Manual Fmoc-SPPS p214
7.2.2.3 *N*-Alloc Protection and Dawson Dbz Resin Loading p215
7.2.2.4 *N*-Alloc Deprotection p215
7.2.2.5 Nbz Formation p215
7.2.2.6 Peptide Cleavage from Acid-Labile Resin p215
7.2.2.7 Reductive Peptide Cleavage from Acid-Labile Resin p216
7.2.3 Syntheses of CCL2 Peptide Fragments p216
7.3 Chemotaxis Assays p219
7.3.1 Materials and Method p219
7.4 References p220

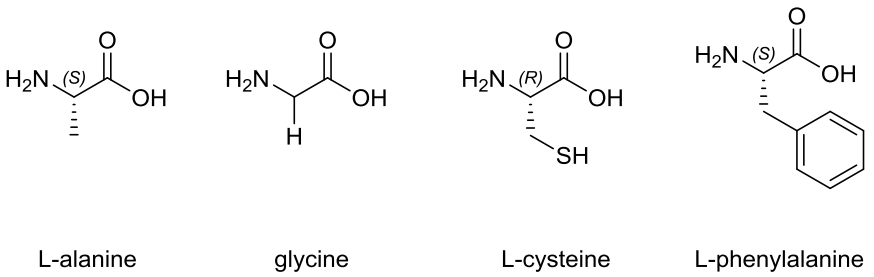
Appendices
Appendix 1: 4-Fluoroproline Synthesis p222
A1.1 The Influence of Fluoroproline p222
A1.2 Synthesis of cis-Fluoroproline p222
A1.3 Incorporation of cis-Fluoroproline into Bioactive Peptides p223
A1.3.1 Incorporation into Lipopeptides p223
A1.3.2 Incorporation into DKP CCL2 Induced Chemotaxis Inhibitors p224
A1.4 Experimental p225
A1.5 References p227
Appendix 2: Ciliatamide B Synthesis p228
A2.1 Anti-Parasitic Lipopeptides: The Ciliatamides p228
A2.2 The Synthesis of Ciliatamide B p229
A2.2 Experimental p231
A2.3 References p233

Chapter 1 : Introduction

1.1 A General Introduction to Proteins

Proteins are naturally occurring macromolecules that perform a wide range of critical cellular functions in all living organisms. Protein functions range from structural and mechanical roles to potent and specific enzyme catalysts. Proteins are polypeptides that consist of a linear sequence of amino acids joined together with multiple peptide bonds and they can vary in length from tens to thousands of amino acids. The diversity and specificity of a given protein derives from its sequence and folding.

In general, naturally occurring amino acids adopt an L configuration, in which, the α-carbon has the *S* stereochemistry (**Figure 1.1**). The exceptions are cysteine (*R*-configuration) and glycine (no chiral centre).



L-alanine glycine L-cysteine L-phenylalanine

Figure 1.1: Naturally occurring α-amino acids: L-alanine, glycine, L-cysteine and L-phenylalanine.

At the most basic level the structure of a protein is defined by the linear sequence of amino acids: known as its primary structure. The long linear sequence of amino acids can interact with one another and form localised regions of defined secondary structure. The most common secondary structural motifs are the α-helix and the β-sheet (**Figure 1.2**).

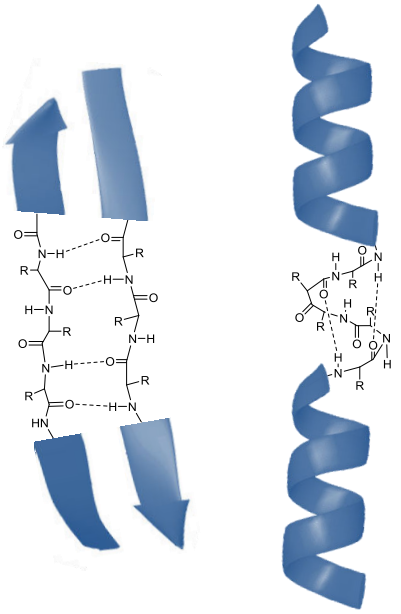


Figure 1.2: Protein secondary structural motifs: ß-sheet (left) and α-Helix (right). Amide hydrogen bonds are indicated by a dashed line (- - -).

In larger proteins a tertiary structure is formed whereby the interactions between secondary structures cause the protein to retain a highly specific shape. The shape adopted allows the protein to function in a very effective and specific manner, e.g. a structurally protein, enzyme or receptor (**Figure 1.3**).[1]

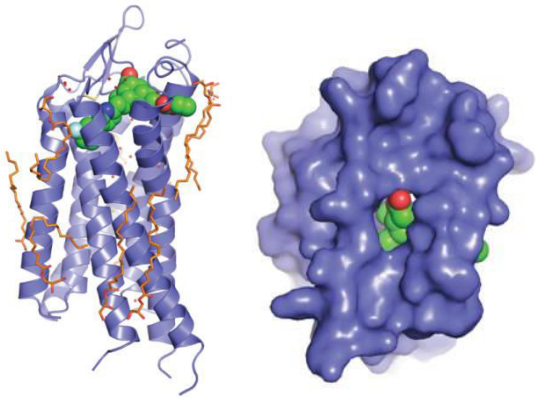


Figure 1.3: The crystal structure of human protease activated receptor 1.[1]

1.2 Chemokines

1.2.1 Introduction to Chemokines

Chemokines (*chemo*attractant cyto*kines*) are a family of small heparin-binding proteins of 8-10kDa (~60-100 amino acids) in size.[2] The first chemokine was discovered in 1977, CXCL4[3] and to date over 45 distinct chemokines (ligands) and 20 chemokine receptors are known.[4] Chemokines are structurally related to cytokines (derived from the Greek: "cyto" meaning cell and "kinos" meaning movement) and they are secreted from many cell types/tissues. Chemokines play a role in directing the inflammatory response by selectively coordinating the migration of immune system cells (monocytes, T-lymphocytes, mast cells, eosinophils and neutrophils)[5] (**Section 1.2.3**).[6]

Chemokine ligands and receptors are targets for anti-inflammatory therapy (**Section 1.4** and **Chapter 4**) due to their key involvement in a variety different inflammatory and autoimmune conditions including multiple sclerosis, psoriasis and rheumatoid arthritis as well as in tissue transplantation and graft versus host disease (GVHD):[4] Where the cells of transplanted tissue are capable of chemokine production and prone to disadvantageous lymphocyte recruitment, host attack and chronic inflammation.[6,7] **Table 1.1** shows the organs affected and the network of chemokine ligands and receptors expressed during the timeline of GVHD.

Organ	Chemokine Receptor Expression	Chemokine Ligand Expression
Lungs	CXCR3, CXCR6, CCR1, CCR2, CCR5, XCR1	CXCL1, CXCL2, CXCL9, CXCL10, CXCL11, CCL2, CCL3, CCL3, CCL4, CCL5, CCL8, CCL12, XCL1
Liver	CXCR2, CXCR3, CXCR6, CCR1, CCR2, CCR5, XCR1	CXCL1, CXCL9, CXCL10, CXCL11, CXCL16, CCL2, CCL3, CCL4, CCL5, CCL12, CCL20, XCL1
Intestines	CXCR3, CXCR6, CCR1, CCR2, CCR5, CCR6, CX_3CR1	CXCL9, CXCL10, CXCL11, CXCL16, CCL2, CCL3, CCL5, CCL20, XCL1, CX_3CL1
Skin	CXCR3, CCR1, CCR2, CCR4, CCR5, CCR10, XCR1	CXCL1, CXCL2, CXCL9, CXCL 10, CXCL11, CCL2, CCL5, CCL6, CCL7, CCL8, CCL9, CCL11, CCL12, CCL17, CCL19, CCL20, CCL27, XCL1
Lymphoid Organs	CXCR3, CCR2, CCR5	CXCL9, CXCL10, CXCL11, CCL2, CCL3, CCL4, CCL5, CCL19, CCL21

Table 1.1: The important organs in relation to GVHD and the various chemokine and chemokine receptors that they express during the timeline of the disease.[8]

Chemokines ligands and receptors have also been implicated in playing a key role in diseases such as: asthma, chronic obstructive pulmonary disease (COPD), HIV/AIDs, types of vascular disease and more recently cancer.[9] A variety of cancer cell types express a non-random set of chemokine receptors and it is thought that the chemokine transport network plays a significant role in metastasis.[10]

1.2.2 Classification of Chemokines

Chemokines are classified in two ways, based on either their (primary) structure similarities or biological function. The primary sequence of amino acids can vary from between 20% to 90%. However, classification of chemokine proteins is based on the conserved cysteine residues in the N-terminal section which leads to four distinct sub-classes with shared structural similarity between family members.[9] For example, CCL2 shares 61% sequence homology with CCL4 and CCL8 and 71% for CCL7.[11] Although members of each class have structural similarities, a diverse chemotactic ability is observed. For example in the CC subset; CCL2, CCL4, CCL5 are chemotactic for monocytes but interact through two different receptors.[5] CCL2 and CCL5 are chemotactic to memory T-cells[12] whereas CCL4 interacts with naïve T-cells.[13] Of the aforementioned chemokines, CCL5 is the only ligand that shows chemotactic activation of eosinophils.[14]

The modern nomenclature of the four chemokine classes describes the subset by its type and relation of N-terminal cysteine residues: XC (or C), CC, CXC and CX_3C (**Table 1.2**).[15]

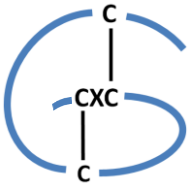
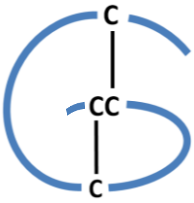
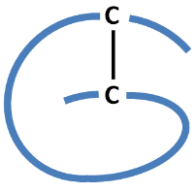
Structure	Type	Class Name	Description
	α	CXC	The first two cysteines (C) are separated by one amino acid (X), and form two disulfide bonds between their 1st and 3rd Cys and between the 2nd and 4th.
	β	CC	The first two cysteines are positioned next to each other, and form two disulfide bonds between their 1st and 3rd Cys and between the 2nd and 4th.
	γ	XC	This class has only one cysteine in the N-terminus and hence only one disulfide bond.

Table 1.2: The simplified structures of 3* major classes/families of chemokine, indicating cysteine residue (C), disulfide bonds (|) and general amino acid (X). *A fourth category of chemokine that only contains only 1 member, Fractalkine (CX_3CL1, 373 AAs), is denoted CX_3C.[16]

Chemokines are further classified based on their physiological features/role into either homeostatic or inflammatory.[17] These features include the cellular distribution of their receptors as well as the location and condition of their production.[17] Homeostatic or "housekeeping" chemokines are involved in immune surveillance and cell migration under the normal processes of tissue maintenance or development e.g. directing cells to secondary lymphoid organs, the bone marrow and thymus during hematopoiesis.[18] Therefore, no source cell stimulation is required for the secretion of homeostatic chemokines.[9] Inflammatory or "inducible" chemokines are expressed by stimulated cells in inflamed tissues and thus, direct immune system cells to the site of injury. Stimulation can be in response to interaction with pro-inflammatory cytokines or pathogens.[4,17]

1.2.3 The Simplified Mode of Action of Inflammatory Chemokines

Generally, inflammatory chemokines are produced locally within a tissue and attract different types of leukocytes to sites of inflammation and infection. The mode of action of chemokines are reasonably homogenous between the different classes/subsets:[19]

Upon injury an up-regulation of selectin and actin (adhesion) molecules on immune system cell surfaces is observed. The binding of these adhesion molecules with their ligands such as fibronectin, collogen and laminin mediates transendothelial chemotaxis (TEC) and causes selectin mediated "rolling" of the leukocytes along the endothelial membrane.[20] The "rolling" decelerates immune system cells as they circulate in the blood stream. Following inflammation/injury chemokines are also released from numerous different cells at the site of injury and they then diffuse into the vascular space creating a concentration gradient. Leukocytes are attracted to higher concentration and thus the origin of chemokine release (site of injury). See **Figure 1.4**.

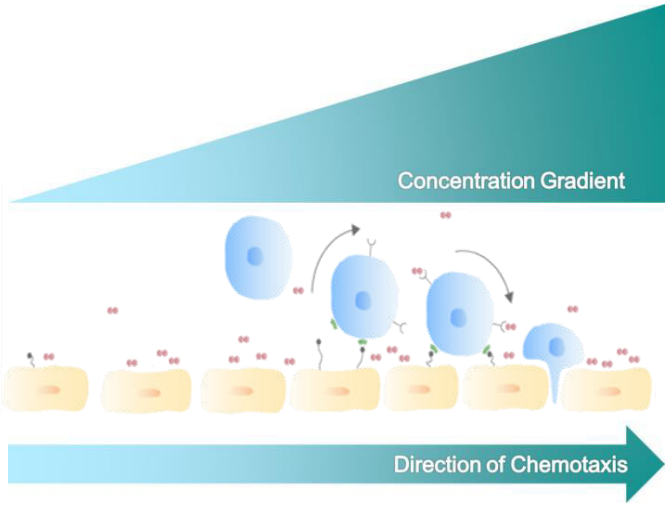


Figure 1.4: Shows the movement of a leukocyte cells (leukocyte extravasation) towards a greater chemokine concentration and thereby to the site of inflammation or infection. At the site of inflammation/infection the moving cell initially rolls along the endothelium surface (mediated by selectin interactions) before adhering more strongly upon chemokine binding. Cell adhesion from integrin/CAM interactions leads to transmigration of the leukocyte through the endothelium.

The concentration gradient is maintained by interactions of chemokines with endothelial cell surface glycosaminoglycans (GAGs), this effectively anchors chemokine proteins to the endothelial cell lining. The selectin mediated "rolling" of leukocyte cells along the lining allows for leukocyte receptors to come into contact and bind with the anchored

chemokines (that are in highest concentration) at the site of injury. This binding is mediated by seven-transmembrane domain (7TM) receptors coupled to guanosine-5'-triphosphate (GTP)-binding proteins able to respond to extracellular molecules, initiate intracellular signal transduction pathways and perform cellular responses. Upon binding of the chemokine ligand to a chemokine receptor a number of drastic and fast effects take place e.g. the stimulation produces appendages on the leukocytes that act to grasp/adhere to endothelial cells before finally migration inside tissues as seen in the **Figure 1.5**.[19]

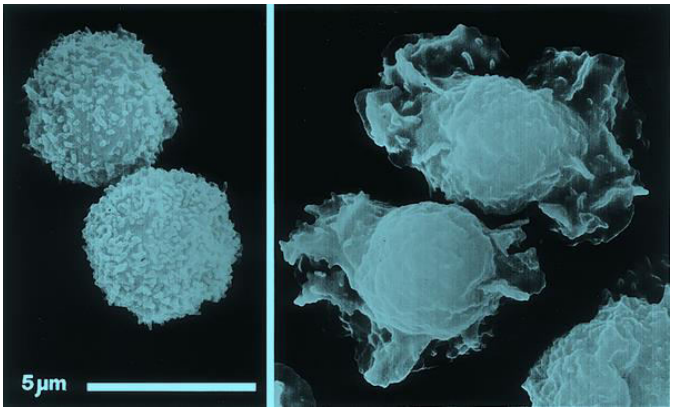


Figure 1.5: The shape change (production of appendages) of leukocyte cells upon chemokine addition.[19]

1.2.4 The Specific and Non-Specific Nature of Chemokines

In general each chemokine family interacts specifically with certain immune system cells. For example, CXC interacts with granulocytes or polymorphonuclear leukocytes (of note: CXCL8 and CXCL10 are members of this subset yet they are also chemotactic to T-lymphocytes). The CC family are chemotactic to mononuclear cells (e.g. monocytes and lymphocytes) and the XC family to lymphocytes.[21] This specificity is granted by the makeup of receptors on the leukocyte cell surface; a list of selected chemokines, alternative names, protein length and the receptor/receptors to which they interact is given in **Table 1.3.**

Chemokine	Functional Name	Protein Length	Chemokine Receptor
CCL1	I-309	74	CCR8
CCL2	MCP-1	76	CCR2
CCL3	MIP-1α	70	CCR1/CCR5
CCL4	MIP-1β	69	CCR5
CCL5	Eotaxin	68	CCR1/CCR3/CCR5
CCL11	RANTES	74	CCR3
CCL17	TARC	71	CCR4
CCL19	MIP-3β / ELC	77	CCR7
CCL20	MIP-3α	70	CCR6
CCL22	MDC	69	CCR4
CCL25	TECK	127	CCR9
CCL26	Eotaxin-3	71	CCR3/CCR10
CXCL8	IL-8	77	CXCR1/CXCR2
CXCL10	IP-10	77	CXCR3
CXCL11	I-TAC	73	CXCR3
CXCL12	SDF-1	68	CXCR4
CXCL13	BCA-1	87	CXCR5
CX3CL1	Fractalkine	76	CX3CR1
XCL1	Lymphotacin	92	XCR1

Table 1.3: Selected chemokines ligands and their known receptors. The inflammatory chemokines shaded.

Chemokine ligand to chemokine receptor interactions can be both specific and non-specific with many chemokine receptors able to bind different ligands with high-affinity. The ligands themselves also show specific and non-specific affinity to receptors leading to the complex set of protein-protein interactions seen in **Figure 1.6**.[22,23] This specific and non-specific nature of binding is a property of both ligand and receptor. However, at present this property is only partially understood. Hence, research into the detailed structural basis of these specific protein-protein interactions is vital.

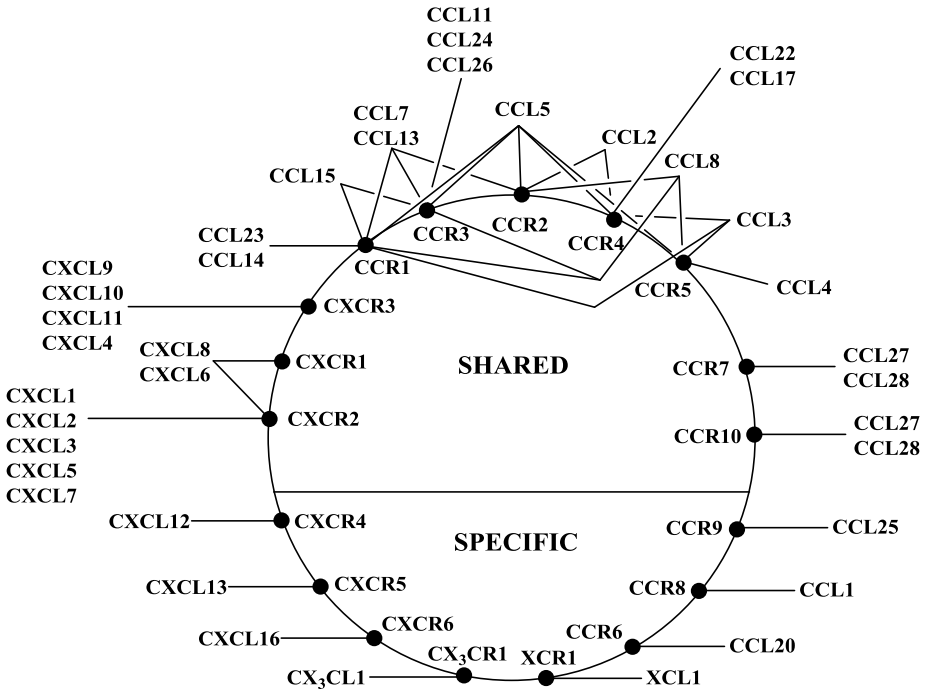
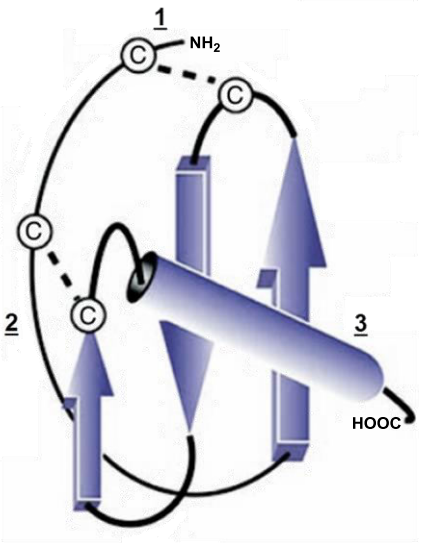


Figure 1.6: The specific and non-specific nature of the relationship between 18 chemokine receptors (inner circle) and 40 ligands (outer circle).[23] The members of a family often competitively bind to the same receptor. However, there is no overlap between families e.g. no CC chemokine interacts with a CXC receptor.[9]

Chemokine ligands can often competitively bind to the same receptor if they are within same family (e.g. C, CC or CXC) and analysis of every known chemokine ligand structure (using NMR studies and crystal structures) has shown a remarkable similarity in their overall three dimensional (tertiary) structure (**Figure 1.7**). This is perhaps surprising due to the considerable differences in primary structure exhibited by chemokines ligands**.**



1. A short N-terminal domain that is highly flexible yet constrained by disulfide bonding between the N-terminal cysteine(s).

2. An irregular loop at the beginning of three antiparallel β-sheets

3. An overlaying α-helix (with respect to the β-sheets).

Figure 1.7: General structure of chemokines with β-sheets indicated as arrows, α-helices as cylinders, cysteines as © and disulfide bonds as a dashed line. This general peptide motif is referred to as a *greek key* or the chemokine fold.[9] The major domains are present throughout chemokines: 1, 2, 3.[23,24]

Modified chemokine ligand analogues have shown that the major interaction sites for binding and subsequent activation of chemokines with their chemokine receptors are contained in the flexible (disordered) N-terminal region. Specifically, the short N-terminal domain (**1, Figure 1.7**) and the exposed backbone loop in-between the 2nd and 3rd cysteine (**2, Figure 1.7**).[25] The loop is thought to be recognised and bound first. This in turn presents the N-terminal domain for function.[25]

CXCL8 analogues can be produced that can still bind the receptor but do not retain function where the N-terminal domain is modified. The deletion or substitution of the terminal sequence: Glu-Leu-Arg (ELR) leads to potent antagonists.[26,27] This is also visualised *in vivo* as a chemokines N-terminal domain is susceptible to cleavage *via* proteolytic digestion. In the late stages of inflammation macrophage specific matrix metalloproteinase-12 (MMP-12) inactivates CXC type chemokines *via* cleavage of the aforementioned ELR to form receptor antagonists. Macrophage recruitment is then inhibited and the cell returns to normal structure and function.[28]

The complete scope of modified analogues of CCL2 and their significance to the mechanism of action of CCL2 is shown in **Table 1.4**.

$Q_1PD_3AI_5N_6AP_8V_9T_{10}CCY_{13}NFTNR_{18}K_{19}ISVQR_{24}LASY_{28}RR_{30}ITSSK_{35}CPK_{38}EAVIFKTIVAK_{49}EICADPKQK_{58}WVQDSMDH_{66}LD_{68}KQTQTPKT$

Residue	Binding	Change	Outcome	Ref
Gln/Glp* - Q1 or <E1		Removal	-reduced *in vitro* activity	[29]
PDAINAP		Removal	-reduced CRR2 binding -reduced signalling -reduced *in vitro* activity	[29,32]
DAIN		D/I/N to A	-reduced *in vitro* activity	[32]
Pro – P8	dimer	P to A	-loss of dimerization -retained *in vitro* activity -no *in vivo* activity	[31,32,35]
Val – V9		V to A	-slightly reduced activity	[32]
Thr – T10		T to R	-reduced activity -reduced CCR2 binding	[36]
Tyr – Y13	CCR2 dimer signalling	Y to A/I	-no in vitro activity -reduced dimerization -slight agonist activity**	[35,36]
Tyr – Y13	CCR2 dimer signalling	Y to F	-unaffected CCR2 binding -unaffected *in vitro* activity	[34]
Tyr – Y13	CCR2 dimer signalling	Y to W/L/H	-reduced *in vitro* activity -agonist activity** -reduced CCR2 binding	[34]
Arg – R18	GAG	K to A	-reduced GAG binding -unaffected *in vitro* activity -no *in vivo* activity	[30,31]
Lys – K19	GAG	K to A	-reduced GAG binding -unaffected *in vitro* activity -no *in vivo* activity	[30,31]
Arg – R24	CCR2 GAG	R to A	-reduced CCR2 binding -reduced GAG binding	[30,33,34]
Tyr – Y28		Y to D	-reduced *in vitro* activity.	[33]
Arg – R30		R to V/ L	-R30V retained activity -R30L reduced activity	[33]
Lys – K35	dimer	SPK insertion K to E	-no *in vitro* activity -reduced CCR2 binding	[34,36]
Lys – K38	dimer	K to E	-reduced CCR2 binding	[34]
Lys – K49	CCR2 GAG	K to A	-reduced CCR2 binding -reduced GAG binding	[30,34]
Lys – K58	GAG	K to A	-reduced GAG binding -reduced *in vivo* activity	[30]
His – H66	GAG	H to A	-reduced GAG binding -reduced *in vivo* activity	[30]
Asp – D68		D to L	-reduced *in vitro* activity.	[33]

Table 1.4: CCL2 primary sequence and residue profiling. *Glutamine (Q1) (The N-terminal amino acid of CCL2) is converted modified to pyroglutmatic acid (<E) *via* a posttranslational modification. The pyroglutamic acid form offers increased *in vivo* stability against degradation.[37] **Agonist activity: Stimulates cytosolic Ca^{2+} influx and inhibits adenylyl cyclase.

Modified analogues have enabled the sites of various chemokine ligand interactions to be found e.g. with GAGs or other chemokine ligands.

In relation to CCL2, it is known that the "positively" charged residues: Arg18, Lys19, Arg24 Lys38, Lys49 and His66 are important in GAG binding[30] (**Figure 1.8**) and the Pro8 residue in the N-terminal section is vital to dimerization.[35] However, mutant analysis has shown failings in translating *in vitro* activity to *in vivo* activity which is exemplified by the Pro8 to Ala8 mutation. The Pro8 to Ala8 mutant retained *in vitro* activity yet showed an inability to form homodimers.[35] The retention of activity did not translate into an *in vivo* model, emphasising the importance of dimerization in a more complex system.

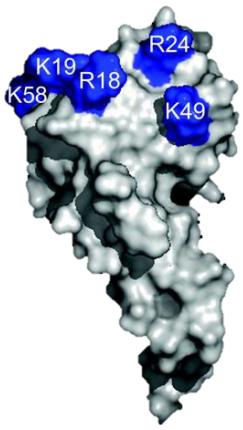


Figure 1.8: GAG binding residues Arg18, Lys19, Arg 24 and Lys49 (blue) mapped on the monomer unit of CCL2. Structure generated by PyMol.[30]

Chemokine ligands gain an additional level of vast complexity from the various and complex array of quaternary structures they can form: such as heterodimers, homodimers and oligomers[38] and in many cases these structures are known to be critical to *in vivo* function.[39,40] Two distinct dimerization motifs and they are attributed to the two major families, CC and CXC.[40,41] CXC dimers share a motif that is formed by the interactions of amino acids in the first ß-sheet strand and gives rise to a six-stranded ß-sheet overlapped by two α-helices.[42] CC dimers form an elongated structure using interactions between residues in the N-termini (**Figure 1.9**).[43] The CC chemokine: CCL20 is known to form a "CXC-like" dimer.[44]

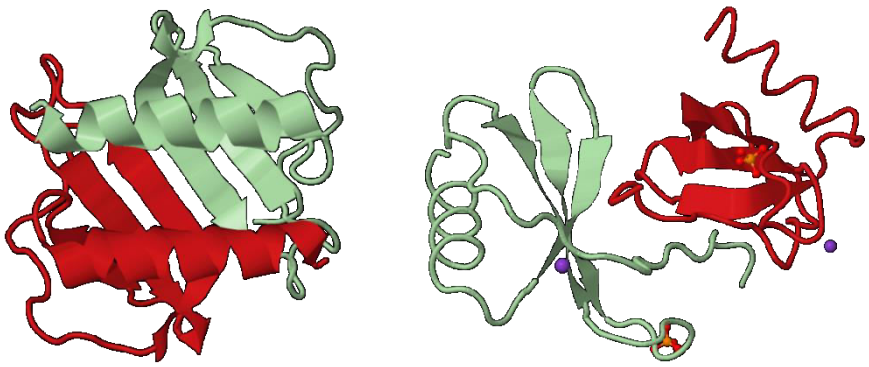


Figure 1.9: The homodimeric structures of CXCL8 (left) and CCL2 (right). The CXCL8 structure was solved by X-ray crystallograply (2.00 Å).[45] The CCL2 structure was solved by X-ray crystallograply (1.90 Å).[43] Subunits are indicated.

As well as ligand dimerization, over half of chemokine receptors are known to exist in heterodimer-, homodimer- or oligo-merization motifs. The varied motifs alter the properties of the receptor (when compared to its monomeric form)[46] in agonist activation, antagonist inhibition, G-protein coupling/signaling and internalization/desensitization of GPCRs.[46] However, the stimulation and the consequences of stimulation of these dimers are poorly understood. Extensive research on CCR5 and CCR2 hetero- and homo-dimers has not led to a consensus on either front. Some evidence suggests that stimulation of receptor-dimer formation requires a ligand interaction[47-49] while other evidence suggests dimerization is ligand independent.[49-52] In terms of activity, E-Asmar *et al.* found trans-inhibition by the ligand to the opposite receptor partner to be a competitive cross-inhibition.[51] Mellado *et al.* found that CCR5/CCR2 heterodimers mediated cell adhesion but not chemotaxis and activity required lower chemokine concentrations.[49] The difficulties in forming an extensive description even in this simplified 2-receptor system perhaps illustrate the challenge of in depth analysis of chemokine ligand/receptor interactions as a whole.

1.3 Chemokine Mediated Immune Response

1.3.1 Inflammation

Inflammation is a complex immunovascular response that heavily involves the chemokine directed movement of immune system cells to damaging stimuli. The

purpose of inflammation is to eliminate the cause of cell injury, prevent the spread of infectious agents, remove necrotic/damaged cells and pathogens as well as initiate tissue repair processes. When inflammation is a short-term, positive, self-limiting, protective response it is termed *acute* inflammation.[53]

1.3.2 Chronic Inflammation

In certain circumstances the inflammatory process becomes continuous and *chronic*. Chronic inflammation is known to contribute a number of diseases e.g. cancer,[54] heart disease,[55] kidney disease,[56] bone disease,[57] diabetes,[58] pancreatitis[59] and anemia.[60] Aging, poor diet, smoking, low-levels of sex hormones, sleep disorders, physical stress and emotional stress are known as external factors that trigger chronic inflammation.

A variety of opportunities for problematic cellular stress and chronic inflammation arise during organ transplantation (from the donor brain death to host reperfusion) as well as throughout the lifetime of the graft. Chemokines; CCL2, CXCL8 CXCL9, CXCL10 and CXCL11 and their corresponding receptors have been implicated at different stages of transplantation.[61] The production of donor and host chemokines and the inflammatory response are an issue that contributes to the rejection of allografts and host death e.g. fatal ischemia - reperfusion injury.[62]

1.3.3 Oxidative Stress

The most significant effect of chronic inflammation is associated with the excessive production of free radicals such as reactive oxygen (ROS) and nitrogen species (RNS) and the depletion of their scavengers (anti-oxidants).[63] This imbalance is termed oxidative stress.[64] The excess of reactive species cause biological damage that alters the normal function of lipids, DNA and proteins, notably chemokines.[65] However, under normal physiological conditions the low-level production of ROS or RNS is beneficial and acts as defence against pathogens. E.g. in macrophages superoxide ($O_2^{\cdot -}$ from NADH oxidase) and nitric oxide (NO from nitric oxide synthase) help kill invading microbes in phagocytes.[66]

It is known that chemokine activity is significantly affected by post-translational modifications and that chemokines are up-regulated and play an important role in areas of oxidative stress.[67] Specifically, CCL2 (and other chemokines) directing leukocytes to the site of allografts.[68] The inflammatory response and up-regulation of chemokines is known to have significant effects in transplantation and transplantation rejection.

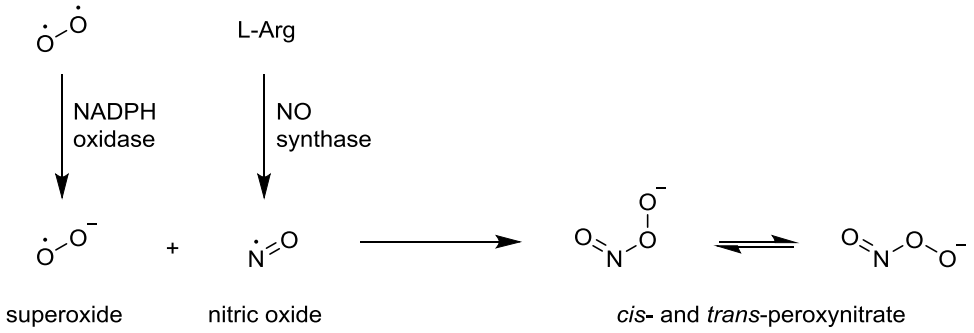
The sources of ROS/RNS molecules range from mitochondrial electron leakage, preioxisomes, NADPH oxidase, amino oxidase, nitric oxide synthase. The types of reactive species that are produced depend on their concentration (of other reactive species), cell environment and cell type. **Table 1.5** shows the structure, source and outcome of selected ROS/RNS species.

Name	Structure	Source	Outcome
Superoxide	$O_2^{\cdot-}$	-enzymatic processes -autooxidation -nonenzymatic electron transfer. (Predominately produced in mitochondria).	$O_2^{\cdot-}$ is a source of other reactive species and is able to act as an oxidising agent (Fe complexes) and a reducing agent (ascorbic acid and tocopherol). However, $O_2^{\cdot-}$ has low intrinsic bioreactivity.
Hydrogen peroxide	H_2O_2	-dismutation reaction (catalyzed *via* superoxide dismutase)	H_2O_2 can deactivate enzymes (e.g. glyceraldehhyde-3-phosphate dehydrogenase) and forms $HO^{\cdot}$ In the presence of transition metals.
Hydroxide radical	$HO^{\cdot}$	-Fenton reaction (H_2O_2 In the prescence of transition metals: Fe^{2+} and Cu^{+}). -Haber-Weiss reaction ($O_2^{\cdot-}$and H_2O_2)	Highly reactive species that affects DNA, proteins, lipids, and carbohydrates. e.g. oxidation of Phe to Tyr.
Nitric oxide	$^{\cdot}NO$	Nitric oxide synthases (conversion of L-Arg to L-Cit)	Regulates enzyme activity (e.g. guanylate cyclase and protein kinases stimulation) and cellular redox processes. Mediates vascular responses (e.g. blood vessel smooth muscle relaxation).
Peroxynitrate	$OONO^{-}$	The rapid reaction of $O_2^{\cdot-}$ and $^{\cdot}NO$	Oxidises DNA, proteins, lipids reacts with CO_2 to form $ONOOCO_2^{-}$ and is protonated to form OONOH.
Peroxynitrous acid	OONOH	Protonation of $OONO^{-}$	Oxidising and nitrating agent that affects proteins, lipids and DNA. Undergoes homolytic fission yielding $^{\cdot}NO_2$ and $HO^{\cdot}$ radicals.

Table 1.5: Selected ROSs and RNSs.[69,70]

1.3.4 Peroxynitrate

Of the ROS/RNS in **Table 1.5** our focus is on the role of peroxynitrate (**Scheme 1.1**) in inflammation and the possible post translational modifications of human chemokine CCL2. Peroxynitrate is a short-lived species (~10 ms at physiological pH) and is therefore, only detected indirectly through the presence of the biomarker: 3-nitrotyrosine. Peroxynitrate is formed in a diffusion-controlled reaction between $NO^{\cdot}$ and $O_2^{\cdot -}$ at a rate of ~1×10^{10} $M^{-1}s^{-1}$ and due to the restricted membrane diffusion and short lifetime of $O_2^{\cdot -}$ is often found at sites of $O_2^{\cdot}$ production.[71] e.g. NADPH oxidases in the plasma membrane or in mitochondrial respiratory complexes.[71] During the inflammatory processes, cells produce high levels of $NO^{\cdot}$ and $O_2^{\cdot -}$ and in turn increased levels of peroxynitrate.



Scheme 1.1: An *in vivo* route to the production of the reactive species: peroxynitrate.

Peroxynitrate exists as a complex set of ionic and radical species (e.g. peroxynitrous acid and nitrosoperoxycarbonate) at different pH and concentrations. It is a potent and versatile oxidizing and nitrating agent[72*] able to modify bioactive molecules (e.g. sulfydryls, ascorbate, DNA, lipids and proteins) and effect pathways (blocking/activation), cell signalling, feedback cycles and ion balance.[73,74]

In relation to protein chemistry, the residues that are susceptible to peroxynitrate posttranslational modification in the form of oxidation and/or nitration are: Tyr, Trp, Phe, His, Cys and Met. **Figure 1.10** shows selected examples of modified amino acids. For a review on this topic see: "Peroxynitrite reactivity with amino acids and proteins.[75]

[*]Gunaydin and Houk modelled (*via* CBS-QB3 calculations on a para-methyphenol system) the mechanisms by which nitration of tyrosine occurs. Their evidence supports a stepwise radical mechanism where the rate limiting step is decomposition of peroxynitrous acid or nitrosoperoxycarbonate.

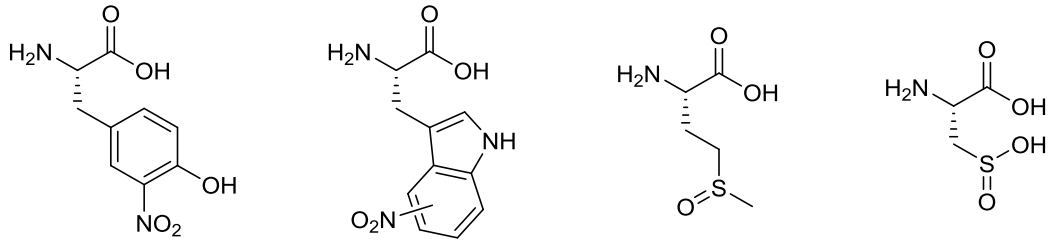


Figure 1.10: Selected protein residue posttranslational modifications carried out by transient peroxynitrate anions/radicals.

The peroxynitrate mediated nitration of tyrosine is a common posttranslational modification and it significantly alters the steric and electronic (hydrogen bonding character, charged state and π-stacking) properties of the residue.[76] When compared to tyrosine, drastic changes are observed (**Figure 1.11**): lowering phenol pK_a (in water: 10.3 to 7.3),[76] increasing spatial volume (ring surface area: ~30Å^2 to ~50Å^2),[77] increasing hydrophobicity[78] and altering redox potential.[79]

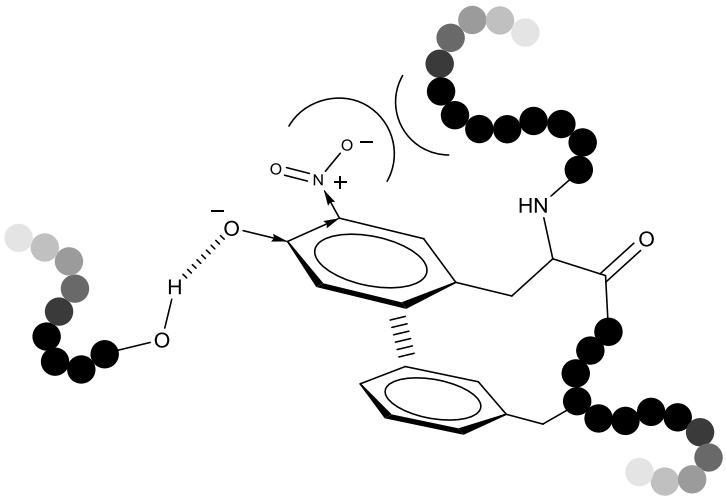


Figure 1.11: A tyrosine residue in a polypeptide. Charged states (+/-), inductive effects (→), steric effects (∩), hydrogen bonding (≡) and π-stacking (≡) are indicated.

Peroxynitrate is also implicated to have a progressive role in transplantation rejection: MacMillan-Crow *et al.* showed that significant peroxynitrate mediated tyrosine nitration of manganese superoxide dismutase (a mitochondrial anti-oxidant enzyme) occurs in human renal allografts.[80] Whereby, peroxynitrate mediated tyrosine nitration deactivates

the anti-oxidant enzyme and this was inferred to have negative extenuating effect leading to increased concentrations of reactive species and in turn, irreversible injury.

1.3.5 Nitration of Chemokine: CCL2

The chemokine: CCL2 is up-regulated at areas of oxidative stress.[67] Therefore, the effect of posttranslational modifications by the reactive nitrogen species: peroxynitrate on the function of CCL2 is of interest.

In model systems, CCL2[81,82] and other chemokines (CCL5[81]) are known to undergo peroxynitrate mediated modifications (in the form of nitration) that abrogate their *in vitro* function (**Figure 1.12**). Sato *et al.* showed that increasing peroxynitrate concentration lowered monocyte chemotactic activity, although Tyr13 nitration was implicated as the cause, no residue analysis was undertaken.[81] In a similar study Molon *et al.* showed that incubation of CCL2 with peroxynitrate followed by MSMS analysis resulted in identification of a nitrated-Trp59 species and that nitrated CCL2 had diminished activity.[82]

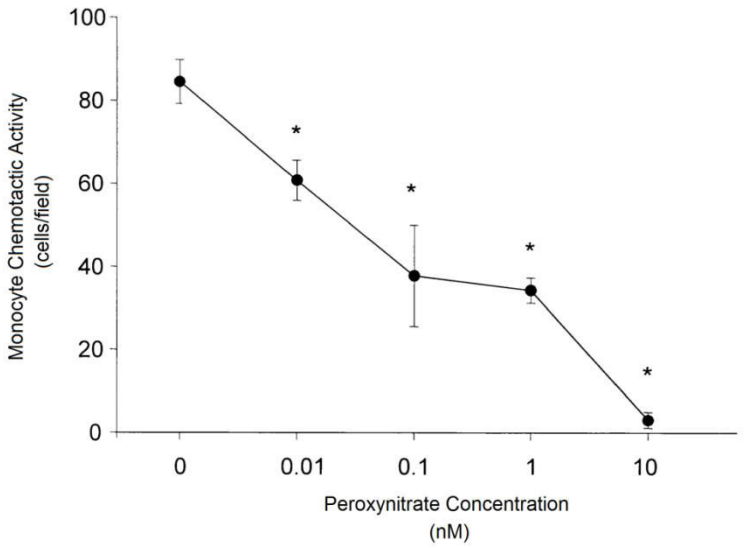


Figure 1.12: Dose-response inhibition of CCL2 mediated monocyte chemotactic activity by peroxynitrate.[81]

Following the aforementioned nitration studies, Dr. C. Barker at Newcastle University recently showed that basic aqueous peroxynitrate was able to nitrate the native chemokine CCL2 in a PBS buffer or H_2O. Mass spectra analysis indicated the presence

of triply nitrated products specifically at Tyr13, Tyr28 and/or Trp59, each nitration giving a mass increase of 45 Da (**Figure 1.13**).

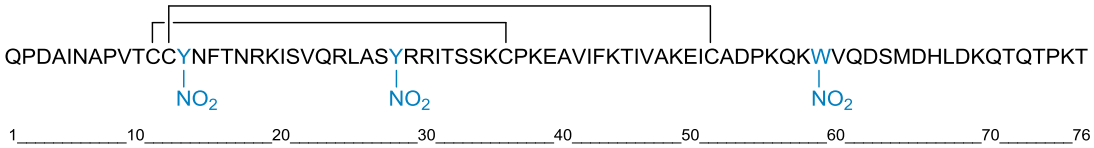


Figure 1.13: Nitrated Tyr13, Tyr28 and Trp59 residues in native human CCL2 (written from N- to C- terminus). Disulfide bonds between C11 - C36 and C12 – C52 indicated.

In agreement with Molon *et al.*[82] and Sato *et al.*[81] the nitrated CCL2 was shown to have decreased biological activity. Nitration of CCL2 decreased CCR2 binding ability and the ability of CCL2 to recruit monocytes in a bare-membrane chemotaxis assay. The nitration of CCL2 also abrogated its ability to interact with GAGs, such as heparin sulfate. GAG binding is vital to *in vivo* function as it allows chemokines to concentrate at the site of inflammation and thus, is crucial for the recruitment of leukocytes. Therefore, in both *in vitro* and *in vivo* chemotaxis assays requiring GAG binding, leukocyte recruitment was prevented. Although much more work is needed, these findings suggest that CCL2 nitration, a modification which could occur during inflammation and oxidative stress could potentially be an important *in vivo* anti-inflammatory mechanism.

However, in this study the nitrated CCL2 was created by rapid reaction with peroxynitrite and the reaction did not go to completion. The chemokine used for the biological assays was therefore a heterogenous mix of protein, with one, two or three of the potential nitration sites being modified on each protein. In order to determine the importance of each of the three residues in the loss of CCL2 biological activity, site-specific synthesis of individually nitrated CCL2 is needed (**Chapter 5**).

1.4 Chemokines as Biological Targets

1.4.1 Therapeutics to Target Chemokine/Chemokine Receptor Interactions

Chemokine/Chemokine receptor interactions are new targets for therapeutics due to their aforementioned role in a variety of conditions: anti-inflammation, growth regulation, HIV-1 infection, embryologic development and angiogenesis (growth of new blood vessels from existing ones).[83] **Table 1.6** shows selected examples of the roles of chemokines in clinical conditions.[84]

Type	Disease	Chemokine	Chemokine Receptor
Inflammatory	Arthritis; colitis	CCR, CXCR	Inflammatory CC, CXC
Infectious	Bacterial/viral Infection; sepsis	CXCR4, CCR5	Inflammatory CC, CXC
Autoimmune	Rheumatoid arthritis, systemic lupus erythematosus	CCR	Inflammatory CC
Allergic	Asthma	CCR3, CCR4, CCR8	CCL11, CCL22, CCL1
Neoplasia	Metastasis; angiogenesis	CXCR4, CCR7	CXCL12, CCL19
Graft rejection	Heart, kidney, lung allograft	CCR5, CXCR1, CXCR3	CCL2, CCL3, CCL5, CXCL8, CXCL9, CXCL10, CXCL11
Vascular	Atherosclerosis; ischemia-reperfusion	CCR2, CXC_3R1	CCL2, CX_3CL1

Table 1.6: The involvement of chemokine and chemokine receptors in various deseases.[84]

Table 1.6 shows data compiled in 2005 and at the time, 7 chemokine receptors (CCR1, CCR2, CCR3, CCR5, CXCR1, CXCR3 and CXCR4) had at least one inhibitor undergoing clinical trials to target them.[84] In recent years chemokines have become successful pharmaceutical drugs.

Maraviroc (Pfizer) an antiretroviral and Plerixafor an immunostimulant (Genzyme) were granted FDA approval to target the chemokine receptors CCR5 and CXCR4 respectively (**Figure 1.14**).[9,85,86]

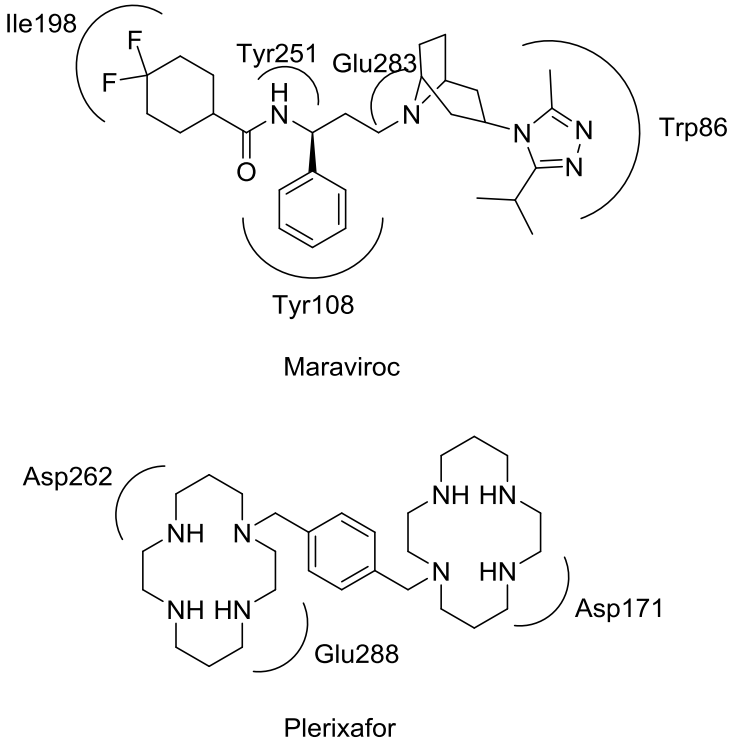


Figure 1.14: CCR5 and CXCR4 inhibitors and registered drugs: Maraviroc and Plerixafor, respectively.[9,87] Presumed bindings between functional groups and the respective receptor residues on are shown.

Small molecule chemokine inhibitors are also an important tool to understand the mechanism of action/binding in these complex protein-protein interaction networks. E.g. UCB35625 showed nanomolar activity (IC_{50} 9.6 nM and 92.7 nM for CCR1-MIP1α and CCR3-eotaxin respectively) as an antagonist to eosinophil chemokine receptors (CCR1 and CCR3) *in vivo*.[88] Although UCB35625 has not fulfilled its potential as an asthma (an eosinophil-mediated inflammatory disorder) therapeutic, it has provided evidence towards elucidating a multi-subsite mechanism of selectivity and action.[88]

The antagonists of chemokine receptors are an important field of medicinal chemistry and knowledge gained is beneficial (**Section 1.5** and **Chapter 3**).[83] Moreover, research in small molecule inhibitors has resulted in numerous compounds reaching clinical trials.[89]

1.4.2 Small Molecule Chemokine Inhibitors: Acylamino-lactams

Oligopeptides (peptide segments) from monocyte chemoattractant protein-1 (CCL2) are known to inhibit chemokine induced cell migration. **Figure 1.15** shows the full primary sequence of CCL2 and the most active peptide segment found in an *in vitro* study by Grainger in 1999.[90] The segment indicated (EICADPKQKWVQ) showed broad inhibition but little selectivity between the CC and CXC families of chemokine with ED_{50} values ranging from 8-14 μM.[90]

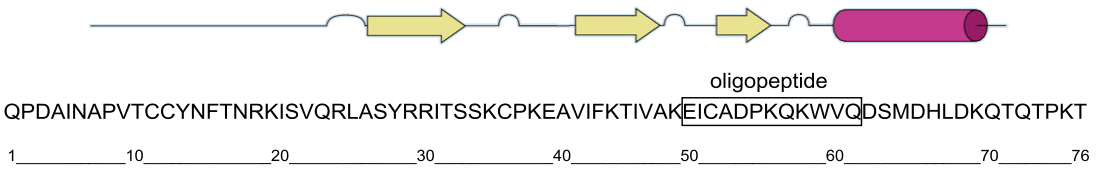
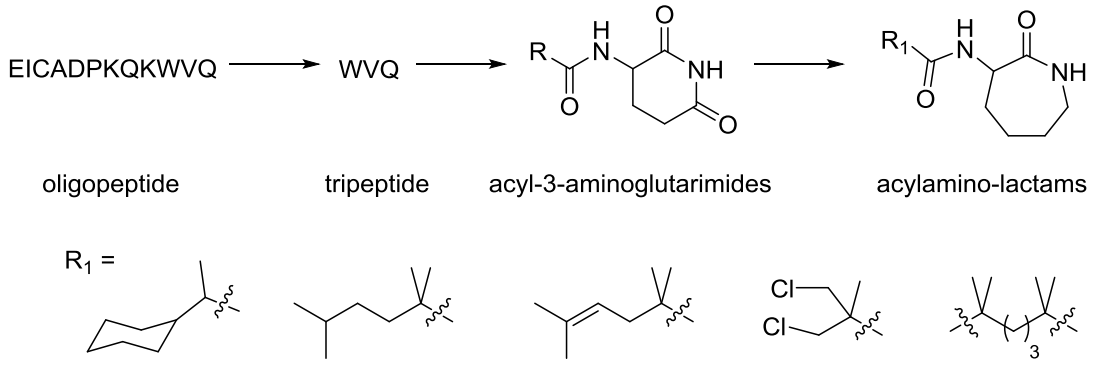


Figure 1.15: The primary structure of CCL2. ß-sheets (arrow) and α-helices (cylinder) are indicated.[43] The 12-residue peptide segment outlined was synthesised *via* SPPS.[90]

The Fox group have synthesised numerous broad-spectrum chemokine inhibitors[91] with research originating from the peptide fragment mentioned above. As the oligopeptides have poor bioavailability and short plasma lifetimes, a structure activity relationship study was undertaken to elucidate the essential structural motifs and to help develop small molecule inhibitors.[92] The *in vitro* activity of a 3 amino acid segment (WVQ) was found to be as potent as the 12 residue oligopeptide (**Scheme 1.2**).[92]



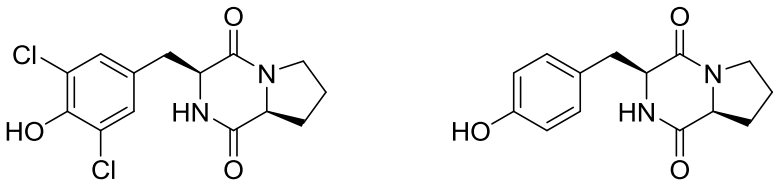
Scheme 1.2: The key structures summarising the work of Fox's group.[93]

A library of compounds originating from this WVQ tripeptide were synthesised and the acyl-3-aminoglutarimides were found to be the best candidates as small molecule

chemokine inhibitors.[93] Further structural refinement led to inhibitors in the form of branched side-chain acylamino-lactams with much improved potency, stability and oral availability (**Scheme 1.2**). The most potent acylamino-lactam has an ED_{50} of 40pM against CCL2 mediated chemotaxis.[91] For more detail on this topic see "Chemokine Receptor Antagonists" by Pease and Horuk.[83,89]

1.4.3 Diketopiperazine Chemokine Inhibitors from Natural Products

In 2009 another class of chemokine inhibitor was discovered and isolated from biological screening of an extract from the fungus: *Leptoxyphium sp*.[94] The inhibitor found: cyclo(13,15-dichloro-L-Pro-L-Tyr) (**Figure 1.16**) is a diketopiperazine (DKP) containing an unnatural dichloroinated tyrosine amino acid. The DKP was tested for inhibition of CCL2 mediated chemotaxis against the de-chlorinated analogue: cyclo(L-Pro-L-Tyr) (**Figure 1.16**). The structure activity relationship study emphasised the importance of the functional groups on the aromatic ring as a 10 to 20 fold drop in activity was observed in the case of cyclo(L-Pro-L-Tyr).[94]



cyclo(13,15-dichloro-L-Pro-L-Tyr) cyclo(L-Pro-L-Tyr)

Figure 1.16: The chlorinated and non-chlorinated DKP structures.

This led the Cobb group to synthesise a small library of analogues (**Figure 1.17**) that included a significant number of unnatural (yet commercially available) aromatic amino acids in order to try and enhance activity and to gain knowledge of the mode of chemotaxis inhibition. The DKPs were synthesised using solution phase peptide synthesis and high temperature cyclisation (**Chapter 4**). Chemotaxis assays confirmed their inhibitory effect on CCL2 induced chemotaxis in an extension of the work by Klausmeyer *et al.*[94]

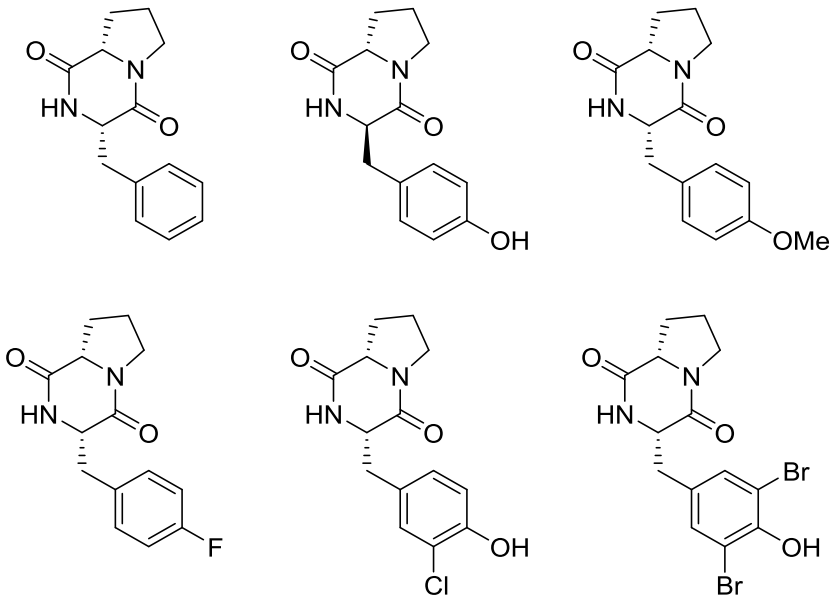


Figure 1.17: The diketopiperazine inhibitors previously synthesised and tested by: Cobb, Ali and Kirby.

1.4.4 Results of the In Vitro Studies of the Chemokine Inhibitors

The results from *in vitro* testing of the DKP library show both specificity and efficacy. A summary of the studies undertaken on the six compounds (**Figure 1.18**) can be visualised below. The bare membrance chemotaxis (BMC) assays measure the inhibition of the compounds against CCL2 induced chemotaxis. Three inhibitors reduced cellular migration by <40% at a concentration of 100 μM (marked with *, **Figure 1.18**).

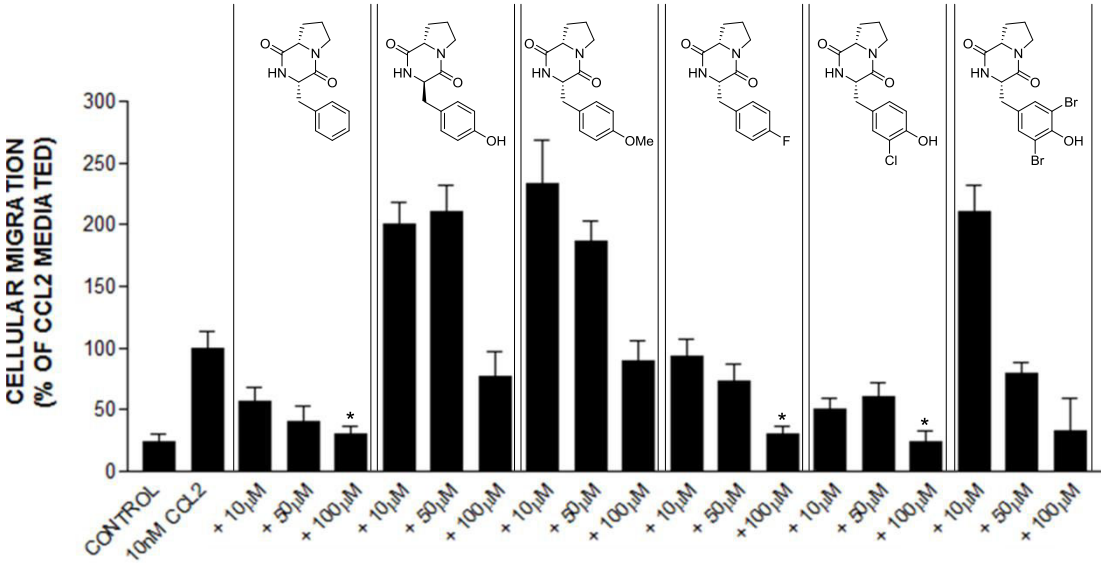


Figure 1.18: Shows the inhibition of all inhibitor compounds at three different concentrations 10, 50 and 100 μM, each with 10nM CCL2. Control contains no CCL2 or inhibitor. The data is normalised to 100% with 10nM CCL2 in the absence of inhibitor

As previously mentioned, chemokine ligand to receptor binding is complicated and a ligand (or ligand complex) can interact with numerous receptors as well as a single receptor binding with various ligands. In relation to CCL2, the ligand-receptor interactions that are known are indicated in **Figure 1.19**. Therefore, To gain information about specificity and the mechanism of inhibition of these DKP molecules a known inhibitor (cyclo(p-fluoro-L-Phe-L-Pro)) was tested against CCL7 and CCL5 (RANTES) induced chemotaxis at 100 μM.

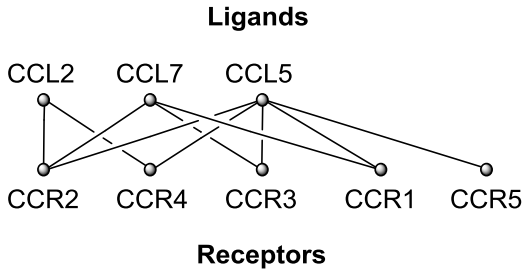


Figure 1.19: Receptor-ligand interaction network for the three chemokines.[23]

CCL7 and CCL5 are known to bind to the CCR2 receptor and induce chemotaxis. However, this process was completely uninhibited in the presence of the DKP (**Figure 1.20**). The compound showed effective *in vitro* inhibition against CCL2 (reduced by

40%, * in **Figure 1.20**), but CCL7 and CCL5 (RANTES) mediated monocyte migration remain uninhibited at a concentration of 100 μM.[24]

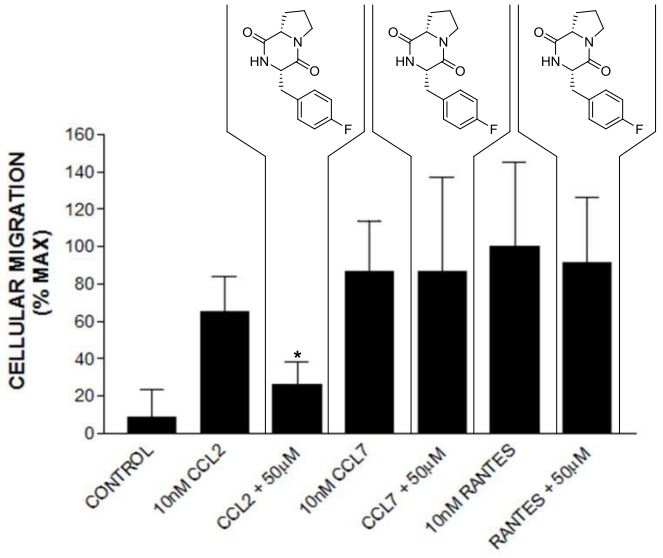


Figure 1.20: The % inhibition of monocyte cell migration by cyclo(p-fluoro-L-Phe-L-Pro) against three different chemokine proteins CCL2, CCL7 and CCL5 (RANTES).

The initial results are promising and they confirm the inhibition of CCL2 induced chemotaxis for the DKP natural product and several analogs as well as offering the posibility of this scaffold to be used as a very specific inhibitor of CCL2 induced chemotaxis without any effect on other CCR2 ligands: CCL7 or CCL5.

Although the activity of all compounds is below what is required for theraupeutic applications, the specificity could enable aspects of unknown chemokine function to be more easily studied by effectively simplifying the complex network of chemokine interactions (**Figure 1.6**) via "knockout" of CCL2 induced chemotaxis.

1.5 Aims

1.5.1 Site-Selective Studies of the Post Translational Nitration of CCL2

The posttranslational nitration modification of CCL2 (outlined in **Section 1.3**) is a biologically relevant route to abrogate the function of CCL2. However, in the aforementioned studies nitrated CCL2 was synthesised *via* a rapid reaction with peroxynitrite (and the reaction did not go to completion). Therefore, the chemokine used

for the biological assays was a heterogenous mix of protein, with one, two, or three of the potential nitration sites being modified on each protein.

Our aim is to determine which nitration site is responsible for loss of function and help study the role of CCL2 in oxidative stress. Thus, a total chemical synthesis (*via* solid phase peptide synthesis (SPPS)) of site-specifically nitrated CCL2 analogues (**Figure 1.21**) is a major goal (**Chapter 3** and **Chapter 5**).

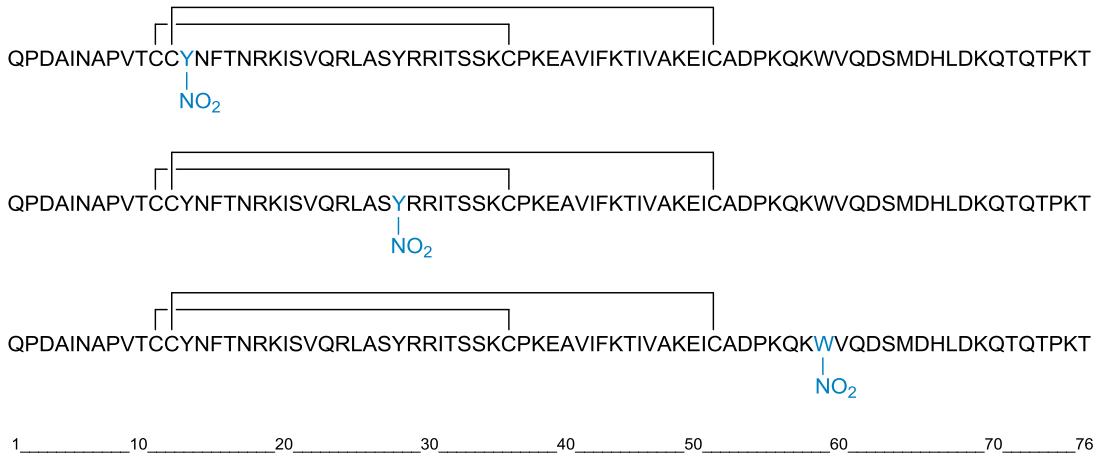
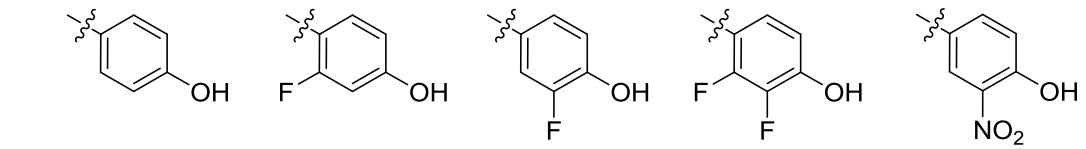


Figure 1.21: Site-specifically mono-nitrated analogues of human chemokine: CCL2

The first step is to develop a total chemical synthesis of native CCL2 and subsequently modify this synthesis to allow incorporation of nitro-aromatic building blocks. Chemotaxis assays would then show which specific nitration has the largest inhibitory effect.

If either of the nitrated tyrosines are found to abrogate function then, to advance this work a series of fluorinated CCL2 analogues would be investigated. This would show the effect of phenolic pK_a modulation (**Figure 1.22**). However, 2-fluorotyrosine and 2,3-difluorotyrosine are not commercially available. Thus, our aim is develop a facile synthetic route to these building blocks (**Chapter 2**).



pK_a = 10.0 9.0 8.4 7.6 7.2-7.5

Figure 1.22: The hydroxyl pK_a of several fluorotyrosine derivatives[95] and 3-nitrotyrosine. The pK_a of 3-nitrotyrosine is known to be significantly affected by the specific peptide environment (e.g. intramolecular π-stacking and hydrogen bonding) and also perturbed due to substrate binding.

Moreover with CCL2 in hand, a series of biophysical studies could be undertaken to help elucidate the mechanism of action (and source of specificity) of the inhibitory DKPs.

1.5.2 Increased Studies on DKP Based CCL2 specific Inhibitors

The selective inhibitors (based on the DKP natural product: cyclo(13,15-dichloro-L-Pro-L-Tyr)) of CCL2 induced chemotaxis described in **Section 1.4** show potential. However, the limited library (**Figure 1.17**) studied provided no clear structure activity relationship and only explored the effect of aromatic ring substiution on activity. Hence, it is our aim is to extend this work by synthesizing and testing a larger library of DKPs with variations in stereochemistry, N-methylation, proline substitution and heteroaromatic substitutents (**Figure 1.23**).

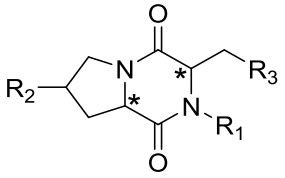


Figure 1.23: A model DKP with variations in stereochemistry (*), N-methylation (R_1), proline substitution (R_2) and heteroaromatic substitutents (R_3).

To do this we envisage optimising the DKP synthetic strategy so that the library can be rapidly expanded. This would enable us to perform the relevant chemotaxis assays to define active structural features and provide more insights into the mode of inhibition (**Chapter 4**). To faciliate this work a number of unnatural amino acids are desired, including heteraromatic (**Chapter 2**) and proline substituted (**Chapter 3**). Therefore, another major aim is develop a facile synthetic route to these building blocks.

1.6. References

1. C. Zhang, Y. Srinivasan, D. H. Arlow, J. J. Fung, D. Palmer, Y. Zheng, H. F. Green, A. Pandey, R. O. Dror, D. E. Shaw, W. I. Weis, S. R. Coughlin and B. K. Kobilka, *Nature*, 2012, **492**, 387-392.
2. P. N. Walsh, R. Biggs and G. Gagnatelli, *Br. J. Haematol.*, 1974, **26**, 405-419.
3. V. Y. Wu, D. A. Walz and L. E. McCoy, *Prep. Biochem.*, 1977, **7**, 479-493.
4. C. Gerard and B. J. Rollins, *Nat. Immunol.*, 2001, **2**, 108-115.
5. S. Sozzani, M. Molino, M. Locati, W. Luini, C. Cerletti, A. Vecchi and A. Mantovani, *J. Immunol.*, 1993, **150**, 1544-1553.
6. K. Vaddi, M. Keller and R. C. Newton, *The Chemokine Facts Book*, Academic Press Harcourt Brace & Company, Publishers. San Diego, 1997.
7. W. W. Hancock, W. Gao, K. L. Faia and V. Csizmadia, *Curr. Opin. Immunol.*, 2000, **12**, 511-516.
8. M. G. M. Castor, V. Pinho and M. M. Teixeira, *Front. Pharmacol.*, 2012, **3**.
9. S. L. Deshmane, S. Kremlev, S. Amini and B. E. Sawaya, *J. Interferon Cytokine Res.*, 2009, **29**, 313-326.
10. A. Zlotnik, *Int. J. Cancer*, 2006, **119**, 2026-2029.
11. E. Van Coillie, J. Van Damme and G. Opdenakker, *Cytokine Growth Factor Rev.*, 1999, **10**, 61-86.
12. M. W. Carr, S. J. Roth, E. Luther, S. S. Rose and T. A. Springer, *Proc. Natl. Acad. Sci.*, 1994, **91**, 3652-3656.
13. D. H. Adams, L. Harvath, D. P. Bottaro, R. Interrante, G. Catalano, Y. Tanaka, A. Strain, S. G. Hubscher and S. Shaw, *Proc. Natl. Acad. Sci.*, 1994, **91**, 7144-7148.
14. M. Ebisawa, T. Yamada, C. Bickel, D. Klunk and R. P. Schleimer, *J. Immunol.*, 1994, **153**, 2153-2160.
15. B. J. Rollins, *Blood*, 1997, **90**, 909-928.
16. H. Umehara, E. T. Bloom, T. Okazaki, Y. Nagano, O. Yoshie and T. Imai, *Arterioscler. Thromb. Vasc. Biol.*, 2004, **24**, 34-40.
17. B. Moser and P. Loetscher, *Nat. Immunol.*, 2001, **2**, 123-128.
18. W. Wagner, C. Roderburg, F. Wein, A. Diehlmann, M. Frankhauser, R. Schubert, V. Eckstein and A. D. Ho, *Stem Cells*, 2007, **25**, 2638-2647.
19. M. Baggiolini, *Nature*, 1998, **392**, 565-568.
20. B. A. Imhof and M. Aurrand-Lions, *Nat. Rev. Immunol.*, 2004, **4**, 432-444.
21. J. J. Oppenheim, C. O. C. Zachariae, N. Mukaida and K. Matsushima, *Annu. Rev. Immunol.*, 1991, **9**, 617-648.
22. M. Philip M, *Cytokine Growth Factor Rev.*, 1996, **7**, 47-64.
23. L. Rajagopalan and K. Rajarathnam, *Biosci. Rep.*, 2006, **26**, 325-339.
24. M. Baggiolini and P. Loetscher, *Immunol. Today*, 2000, **21**, 418-420.
25. I. Clarklewis, K. S. Kim, K. Rajarathnam, J. H. Gong, B. Dewald, B. Moser, M. Baggiolini and B. D. Sykes, *J. Leukoc. Biol.*, 1995, **57**, 703-711.
26. M. Baggiolini, B. Dewald and B. Moser, *Adv. Immunol.*, 1994, **55**, 97-179.
27. M. Baggiolini and I. Clarklewis, *FEBS Lett.*, 1992, **307**, 97-101.
28. R. A. Dean, J. H. Cox, C. L. Bellac, A. Doucet, A. E. Starr and C. M. Overall, *Blood*, 2008, **112**, 3455-3464.
29. J. H. Gong and I. Clark-Lewis, *J. Exp. Med.*, 1995, **181**, 631-640.
30. E. K. Lau, C. D. Paavola, Z. Johnson, J. P. Gaudry, E. Geretti, F. Borlat, A. J. Kungl, A. E. Proudfoot and T. M. Handel, *J. Biol. Chem.*, 2004, **279**, 22294-22305.

31. A. E. I. Proudfoot, T. M. Handel, Z. Johnson, E. K. Lau, P. LiWang, I. Clark-Lewis, F. Borlat, T. N. C. Wells and M. H. Kosco-Vilbois, *Proc Natl. Acad. Sci.*, 2003, **100**, 1885-1890.
32. K. Jarnagin, D. Grunberger, M. Mulkins, B. Wong, S. Hemmerich, C. Paavola, A. Bloom, S. Bhakta, F. Diehl, R. Freedman, D. McCarley, I. Polsky, A. Ping-Tsou, A. Kosaka and T. M. Handel, *Biochemistry*, 1999, **38**, 16167-16177.
33. Y. J. Zhang, B. J. Rutledge and B. J. Rollins, *J. Biol. Chem.*, 1994, **269**, 15918-15924.
34. S. Hemmerich, C. Paavola, A. Bloom, S. Bhakta, R. Freedman, D. Grunberger, J. Krstenansky, S. Lee, D. McCarley, M. Mulkins, B. Wong, J. Pease, L. Mizoue, T. Mirzadegan, I. Polsky, K. Thompson, T. M. Handel and K. Jarnagin, *Biochemistry*, 1999, **38**, 13013-13025.
35. C. D. Paavola, S. Hemmerich, D. Grunberger, I. Polsky, A. Bloom, R. Freedman, M. Mulkins, S. Bhakta, D. McCarley, L. Wiesent, B. Wong, K. Jarnagin and T. M. Handel, *J. Biol. Chem.*, 1998, **273**, 33157-33165.
36. C. J. Beall, S. Mahajan, D. E. Kuhn and P. E. Kolattukudy, *Biochem. J.*, 1996, **313**, 633-640.
37. M. Kruszynski, N. Stowell, A. Das, J. Seideman, P. Tsui, M. Brigham-Burke, J. F. Nemeth, R. Sweet and G. A. Heavner, *J. Pept. Sci.*, 2006, **12**, 25-32.
38. L. G. Czaplewski, J. McKeating, C. J. Craven, L. D. Higgins, V. Appay, A. Brown, T. Dudgeon, L. A. Howard, T. Meyers, J. Owen, S. R. Palan, P. Tan, G. Wilson, N. R. Woods, C. M. Heyworth, B. I. Lord, D. Brotherton, R. Christison, S. Craig, S. Cribbes, R. M. Edwards, S. J. Evans, R. Gilbert, P. Morgan, E. Randle, N. Schofield, P. G. Varley, J. Fisher, J. P. Waltho and M. G. Hunter, *J. Biol. Chem.*, 1999, **274**, 16077-16084.
39. J. F. Paolini, D. Willard, T. Consler, M. Luther and M. S. Krangel, *J. Immunol.*, 1994, **153**, 2704-2717.
40. S. E. Crown, Y. H. Yu, M. D. Sweeney, J. A. Leary and T. M. Handel, *J. Biol. Chem.*, 2006, **281**, 25438-25446.
41. E. J. Fernandez and E. Lolis, *Annu. Rev. Pharmacol. Toxicol.*, 2002, **42**, 469-499.
42. G. M. Clore, E. Appella, M. Yamada, K. Matsushima and A. M. Gronenborn, *Biochemistry*, 1990, **29**, 1689-1696.
43. T. L. Grygiel, A. Teplyakov, G. Obmolova, N. Stowell, R. Holland, J. F. Nemeth, S. C. Pomerantz, M. Kruszynski and G. L. Gilliland, *Biopolymers*, 2010, **94**, 350-359.
44. D. M. Hoover, C. Boulegue, D. Yang, J. J. Oppenheim, K. Tucker, W. Lu and J. Lubkowski, *J. Biol. Chem.*, 2002, **277**, 37647-37654.
45. E. T. Baldwin, I. T. Weber, R. St Charles, J. C. Xuan, E. Appella, M. Yamada, K. Matsushima, B. F. Edwards, G. M. Clore, A. M. Gronenborn and et al., *Proc. Natl. Acad. Sci.*, 1991, **88**, 502-506.
46. E. Jacoby, R. Bouhelal, M. Gerspacher and K. Seuwen, *ChemMedChem*, 2006, **1**, 761-782.
47. J. M. Rodriguez-Frade, A. J. Vila-Coro, A. M. de Ana, J. P. Albar, A. C. Martinez and M. Mellado, *Proc. Natl. Acad. Sci.*, 1999, **96**, 3628-3633.
48. A. J. Vila-Coro, M. Mellado, A. Martin de Ana, P. Lucas, G. del Real, A. C. Martinez and J. M. Rodriguez-Frade, *Proc. Natl. Acad. Sci.*, 2000, **97**, 3388-3393.

49. M. Mellado, J. M. Rodriguez-Frade, A. J. Vila-Coro, S. Fernandez, A. Martin de Ana, D. R. Jones, J. L. Toran and A. C. Martinez, *EMBO J.*, 2001, **20**, 2497-2507.
50. M. Benkirane, D. Y. Jin, R. F. Chun, R. A. Koup and K. T. Jeang, *J. Biol. Chem.*, 1997, **272**, 30603-30606.
51. L. El-Asmar, J. Y. Springael, S. Ballet, E. U. Andrieu, G. Vassart and M. Parmentier, *Mol. Pharmacol.*, 2005, **67**, 460-469.
52. H. Issafras, S. Angers, S. Bulenger, C. Blanpain, M. Parmentier, C. Labbe-Jullie, M. Bouvier and S. Marullo, *J. Biol. Chem.*, 2002, **277**, 34666-34673.
53. L. Ferrero-Miliani, O. H. Nielsen, P. S. Andersen and S. E. Girardin, *Clin. Exp. Immunol.*, 2007, **147**, 227-235.
54. J. K. Kundu and Y. J. Surh, *Mutat. Res.*, 2008, **659**, 15-30.
55. T. Singh and A. B. Newman, *Ageing Res. Rev.*, 2011, **10**, 319-329.
56. G. Glorieux, G. Cohen, J. Jankowski and R. Vanholder, *Semin. Dial.*, 2009, **22**, 423-427.
57. J. J. Cao, *J. Orthop. Surg. Res.*, 2011, **6**, 30.
58. J. P. Bastard, M. Maachi, C. Lagathu, M. J. Kim, M. Caron, H. Vidal, J. Capeau and B. Feve, *Eur. Cytokine Netw.*, 2006, **17**, 4-12.
59. R. K. Jha, Q. Ma, H. Sha and M. Palikhe, *Med. Sci. Monit.*, 2009, **15**, Ra147-156.
60. L. Ferrucci, R. D. Semba, J. M. Guralnik, W. B. Ershler, S. Bandinelli, K. V. Patel, K. Sun, R. C. Woodman, N. C. Andrews, R. J. Cotter, T. Ganz, E. Nemeth and D. L. Longo, *Blood*, 2010, **115**, 3810-3816.
61. C. E. Barker, S. Ali, G. O'Boyle and J. A. Kirby, *Immunology*, 2014, **143**, 138-145.
62. H. K. Eltzschig and T. Eckle, *Nat. Med.*, 2011, **17**, 1391-1401.
63. N. Khansari, Y. Shakiba and M. Mahmoudi, *Recent Pat. Inflamm. Allergy Drug Discov.*, 2009, **3**, 73-80.
64. S. Salzano, P. Checconi, E.-M. Hanschmann, C. H. Lillig, L. D. Bowler, P. Chan, D. Vaudry, M. Mengozzi, L. Coppo, S. Sacre, K. R. Atkuri, B. Sahaf, L. A. Herzenberg, L. A. Herzenberg, L. Mullen and P. Ghezzi, *Proc Natl. Acad. Sci.*, 2014, **111**, 12157-12162.
65. B. S. Berlett and E. R. Stadtman, *J. Biol. Chem.*, 1997, **272**, 20313-20316.
66. M. Valko, D. Leibfritz, J. Moncol, M. T. Cronin, M. Mazur and J. Telser, *Int. J. Biochem. Cell Biol.*, 2007, **39**, 44-84.
67. A. Mortier, J. Van Damme and P. Proost, *Pharmacol. Ther.*, 2008, **120**, 197-217.
68. T. el-Sawy, N. M. Fahmy and R. L. Fairchild, *Curr. Opin. Immunol.*, 2002, **14**, 562-568.
69. A. Phaniendra, D. B. Jestadi and L. Periyasamy, *Indian. J. Clin. Biochem.*, 2015, **30**, 11-26.
70. B. Halliwell, *Plant Physiol.*, 2006, **141**, 312-322.
71. C. Szabo, H. Ischiropoulos and R. Radi, *Nat. Rev. Drug Discov.*, 2007, **6**, 662-680.
72. H. Gunaydin and K. N. Houk, *Chem. Res. Toxicol.*, 2009, **22**, 894-898.
73. R. Radi, J. S. Beckman, K. M. Bush and B. A. Freeman, *J. Biol. Chem.*, 1991, **266**, 4244-4250.
74. B. Alvarez, H. Rubbo, M. Kirk, S. Barnes, B. A. Freeman and R. Radi, *Chem. Res. Toxicol.*, 1996, **9**, 390-396.
75. B. Alvarez and R. Radi, *Amino Acids*, 2003, **25**, 295-311.

76. R. Radi, *Acc. Chem. Res.*, 2013, **46**, 550-559.
77. S. N. Savvides, M. Scheiwein, C. C. Bohme, G. E. Arteel, P. A. Karplus, K. Becker and R. H. Schirmer, *J. Biol. Chem.*, 2002, **277**, 2779-2784.
78. J. M. Souza, I. Choi, Q. Chen, M. Weisse, E. Daikhin, M. Yudkoff, M. Obin, J. Ara, J. Horwitz and H. Ischiropoulos, *Arch. Biochem. Biophys.*, 2000, **380**, 360-366.
79. K. Yokoyama, U. Uhlin and J. Stubbe, *J. Am. Chem. Soc.*, 2010, **132**, 8385-8397.
80. L. A. MacMillan-Crow, J. P. Crow, J. D. Kerby, J. S. Beckman and J. A. Thompson, *Proc. Natl. Acad. Sci.*, 1996, **93**, 11853-11858.
81. E. Sato, K. L. Simpson, M. B. Grisham, S. Koyama and R. A. Robbins, *Am. J. Pathol.*, 1999, **155**, 591-598.
82. B. Molon, S. Ugel, F. Del Pozzo, C. Soldani, S. Zilio, D. Avella, A. De Palma, P. Mauri, A. Monegal, M. Rescigno, B. Savino, P. Colombo, N. Jonjic, S. Pecanic, L. Lazzarato, R. Fruttero, A. Gasco, V. Bronte and A. Viola, *J Exp. Med.*, 2011, **208**, 1949-1962.
83. J. E. Pease and R. Horuk, *Expert Opin. Ther. Patents*, 2009, **19**, 39-58.
84. M. Locati, R. Bonecchi and M. M. Corsi, *Am. J. Clin. Pathol.*, 2005, **123**, S82-95.
85. L. R. Wedderburn, N. Robinson, A. Patel, H. Varsani and P. Woo, *Arthritis Rheum.*, 2000, **43**, 765-774.
86. A. E. Proudfoot, C. A. Power and M. K. Schwarz, *Expert Opin. Investig. Drugs.*, 2010, **19**, 345-355.
87. S. P. Fricker, *Transfus. Med. Hemother.*, 2013, **40**, 237-245.
88. I. Sabroe, M. J. Peck, B. J. Van Keulen, A. Jorritsma, G. Simmons, P. R. Clapham, T. J. Williams and J. E. Pease, *J. Biol. Chem.*, 2000, **275**, 25985-25992.
89. J. E. Pease and R. Horuk, *Expert Opin. Ther. Patents*, 2009, **19**, 199-221.
90. J. Reckless and D. J. Grainger, *Biochem. J.*, 1999, **340**, 803-811.
91. D. J. Fox, J. Reckless, H. Lingard, S. Warren and D. J. Grainger, *J. Med. Chem.*, 2009, **52**, 3591-3595.
92. D. J. Fox, J. Reckless, S. G. Warren and D. J. Grainger, *J. Med. Chem.*, 2002, **45**, 360-370.
93. D. J. Fox, J. Reckless, S. M. Wilbert, I. Greig, S. Warren and D. J. Grainger, *J. Med. Chem.*, 2005, **48**, 867-874.
94. P. Klausmeyer, O. M. Z. Howard, S. M. Shipley and T. G. McCoud, *J. Nat. Prod.*, 2009, **72**, 1369-1372.
95. K. Kim and P. A. Cole, *J. Am. Chem. Soc.*, 1998, **120**, 6851-6858.

Chapter 2 : The Synthesis and Applications of Novel Aromatic & Heteroaromatic Amino Acids

2.1 Introduction

2.1.1 Expanding Nature's "Tool-box" of Amino Acids

Novel (or unnatural) amino acids expand nature's "tool-box" and increase the scope of functionality available for the preparation of peptides, proteins and other related biological molecules. This has led to their extensive use in both biology and chemistry, for example to improve the properties of peptide/protein therapeutics or to investigate biological systems (**Table 2.1**).[1,2] Some selected examples of the applications of novel amino acids include mechanistic probes,[3] structure modulation,[4,5] conformational stability,[6-8] protein complex mapping,[9-11] ^{19}F NMR[11,12] and UV Raman probes.[13,14]

Given the aforementioned applications the developments of efficient methods for the syntheses of novel unnatural amino acids is vital. In relation to the work reported herein, novel amino acids are being used to facilitate rational bioactive peptide design (e.g. Chemokine inhibitors: **Section 2.5.4** and **Chapter 4**) and in the full total chemical SPPS of protein analogues (e.g. human chemokine CCL2: **Chapter 5**).

For more detailed review articles on the synthesis of novel amino acids and related topics see: "Recent Advances in the Synthesis of Unnatural α-Amino Acids - An Updated Version"[15] and "Unnatural Amino Acids in Enzymes and Proteins"[16]

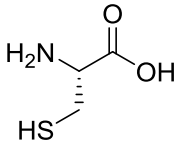
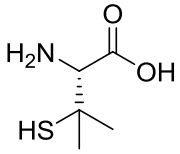
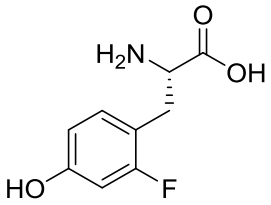
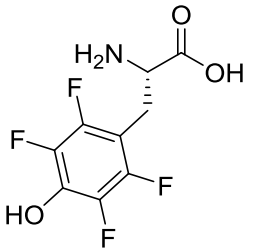
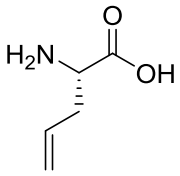
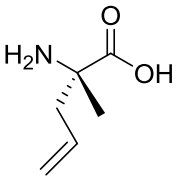
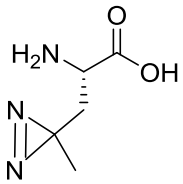
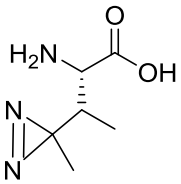
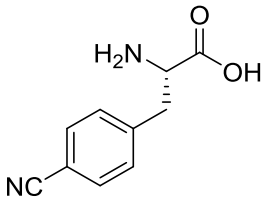
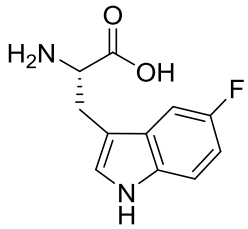
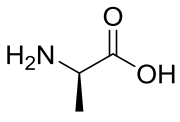
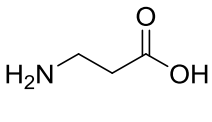
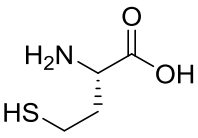
Structure	Biological Relevance	Refs
 Cys  Pen	Utilising steric effects to modulate the coordination number of metal co-factor.	4
 2-fluoro-L-Tyr  2,3,4,5-tetrafluoro-L-Tyr	Tyrosine hydroxyl hydrogen bond modulation *via* introduction of ring fluorines.	5
 L-allyl-Gly  (S)-*a*-allyl-Ala	Structural enforcement to increase peptide stability/activity *via* ring closing metathesis (RCM) cyclisation	6,7
 Photo-Leu  Photo-Ile	Protein-protein interaction mapping *via* photo cross-linking and subsequent digest/MSMS.	10
 4-cyano- L-Phe  5-fluoro-L-Trp	Characteristic spectroscopic (NMR & Raman) probe amino acids that are responsive to the local chemical environment.	12,14
 D-Ala  *B*-Ala  Hcy	To help elucidate the mechanism of action in protein (intien) splicing	3

Table 2.1: Unnatural amino acids and a relevant selected biological use.

2.1.2 Incorporation of Novel/Unnatural Amino Acids into Peptides and Proteins

If the target protein/peptide can be accessed by solid phase peptide syntheses (SPPS) then the site-specific incorporation of any available unnatural amino acid is relatively straightforward. Challenges arise if the amino acid is not commercially available or if the novel amino acid building block (and protection) is unstable to the reaction conditions involved in SPPS. In addition it should be noted that sequential linear SPPS is only feasible for the preparation of peptides of up to 50 amino acids. The upper boundaries are being pushed by new developments in reagents and synthesisers but typically for longer peptides and proteins a ligation approach must be adopted. For a description of the reaction steps, product scope and limitations of ligation strategies in peptide synthesis, see **Chapter 5**.

When the total chemical synthesis of peptides/proteins is not feasible the incorporation of unnatural amino acids may be possible by utilising a biological organism/system. Residue-specific[17] and site-specific methods exist, whereby genetic engineering and thus modifications of a biological organisms synthetic pathways lead to unnatural amino acid incorporation. Certain unnatural amino acids (e.g. norleucine,[17] fluoroproline,[18] and azidohomoalanine,[19] **Figure 2.1**) are *isostructural* to the canonical amino acids and can be tolerated and recognized by the host cells machinery and in turn and incorporated into peptide/proteins. In this biosynthetic route the growth medium of the organism producing the peptide/protein of interest is supplemented with an unnatural amino acid. This strategy has been effectivity utilised in **Section 2.5** to attempt to incorporate the unnatural amino acids: 2-fluorotyrosine, 2,3-difluorotyrosine and *cis*-fluoroproline into lipopeptides from *Bacillus* sp. CS93.[20]

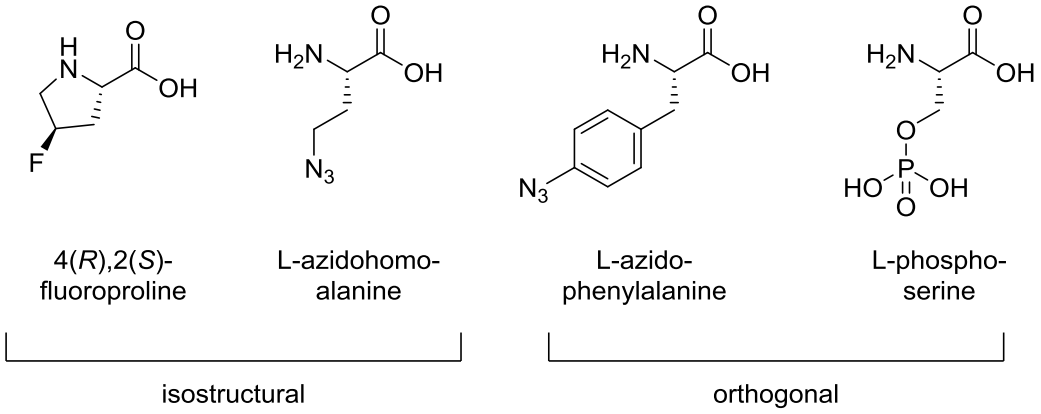


Figure 2.1: Isostructural amino acids: Fluoroproline and azidohomoalanine. Orthogonal amino acids: Azidophenylalanine and phosphoserine.

A significant number of unnatural amino acids (e.g. 4-azidophenylalanine[21] and phosphoserine[22] in **Figure 2.1**) are not tolerated by the host cells translation pathways (*orthogonal*) and require incorporation *via* a more elaborate technique: Stop codon suppression (SCS). SCS utilises orthogonal aminoacyl-tRNA synthetase/tRNA pairs (o-pairs).[23] A variety of o-pair systems exist with the two most proliferant being *Methanocaldococcus jannaschii* TyrRS (mjTyrRS) and *Methanosarcinaceae* PylRSs (PyIRS). mjTyrRS have been responsible for incorporation of around 40 unnatural AAs, mostly orthogonal aromatic derivatives of phenylalanine or tyrosine.[24] PyIRS systems are a recent development but they have already shown remarkable tolerance with respect to both orthogonal aliphatic and aromatic amino acids.[23]

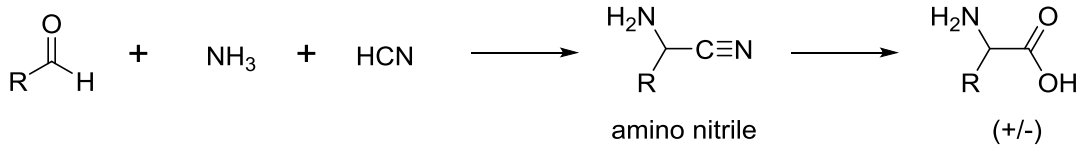
The aforementioned biosynthetic approaches have significant disadvantages. These include; inefficient uptake, toxicity and non-specific inclusion of amino acids that can give unwanted cumulative protein/peptide perturbations.[11] For a recent review of this area see: "Incorporation of unnatural amino acids for synthetic biology"[25] and "Unnatural amino acids in enzymes and proteins".[16]

2.1.3 Previous Syntheses of Unnatural α-Amino Acids

Unnatural amino acids can be categorised by the carbon to which the amine group is attached (i.e. α-, ß-, γ- or δ-) relative to the carboxylic acid. Although, synthetic routes for all of these categories exist, this work will focus on the preparation of α-amino acids.

Biological/enzymatic syntheses and resolutions are the viable routes for the production of natural L-α-amino acids. Unnatural α-amino acids are not always well tolerated by biosynthetic processes as their structure varies from the 20 proteogenic L-amino acids by the side chain functionality and/or the stereochemistry at the α-carbon. Thus, synthetic routes to access optically pure unnatural α-amino acids are a necessity. The synthetic routes fall broadly into three general categories: i) The use of achiral molecules with diastereoisomer/chiral resolution; ii) auxiliaries/catalysts to define the stereochemistry at the α-carbon; and iii) the modification of cheap and available natural amino acids.

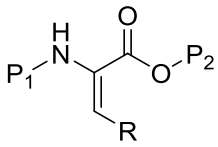
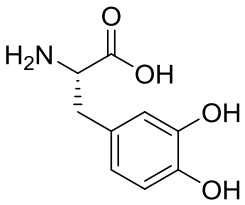
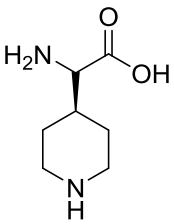
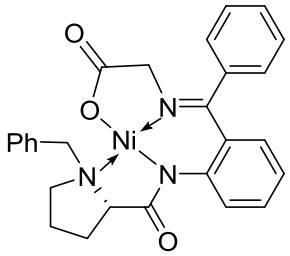
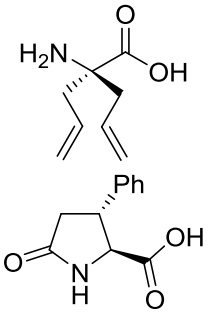
The synthesis of amino acids originates in 1850 when Adolf Strecker developed his three-component reaction (**Scheme 2.1**)[26] and efforts to find high-yielding, atom efficient, asymmetric and functional group tolerant routes has been an important area of research ever since.[15]



Scheme 2.1: The Strecker synthesis. The formation of an amino nitrile from an aldehyde, amine and hydrogen cyanide. The amino nitrile is subsequently hydrolysed to give an amino acid.

Recently, asymmetric versions of the Strecker reaction have been developed using a variety of chiral auxiliaries[27] or chiral catalysts[28] giving a broad product scope with yields and enantiomeric excesses regularly above 90%.[28] However, the reaction is limited as the cyanide nucleophile source (TMSCN) generates volatile and toxic HCN *in-situ*. Recent advances have shown that cheaper and safer alternatives can be used: CNCOOEt[29] and KCN[30] but these strategies elongate the synthesis as an imine (or imine equivalent) must be pre-formed and isolated.

The Strecker reaction is just one of many strategies for α-amino acid syntheses and **Table 2.2** shows three other selected approaches commonly reported in the literature. Entry 1: The asymmetric addition of hydrogen across a pro-chiral α-carbon of a dehydroamino acid. Entry 2: The derivatization of glycine by formation of a nucleophile or electrophile at the α-carbon centre and then the subsequent asymmetric R-group addition. Entry 3: Palladium catalysed cross-coupling reactions (Suzuki,[31] Negishi,[32,33] and Heck[34]) can access libraries of unsaturated amino acids from single chiral amino acid starting materials (e.g. serine[32] and methionine[34]). The Cobb group has focused on developing the synthesis of unnatural amino acids *via* Pd-catalysed cross-coupling, specifically, the Negishi reaction.

Entry	Reaction	Key Intermediate(s)	Example Products	Refs
1	Asymmetric hydrogenation of dehydroamino acids		 	35-37
2	Nucleophilic glycine derivatization (Ni(II) complex Schiff base)			38-40

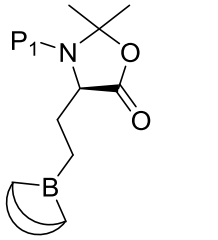
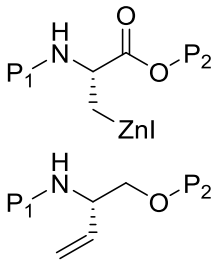
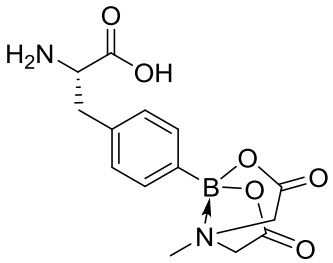
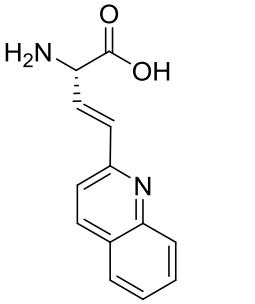
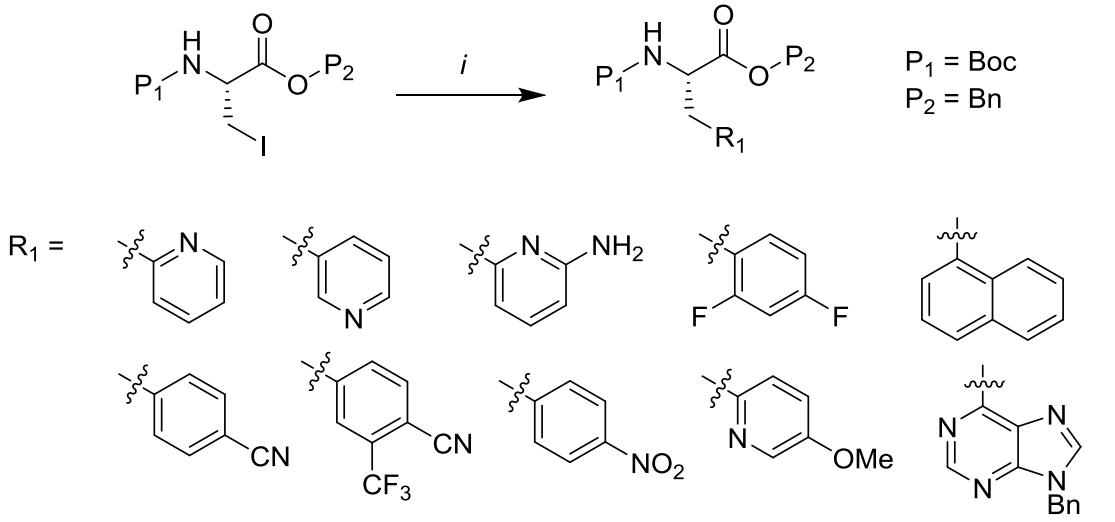
3	Pd-catalysed cross-coupling	 	 	31,33, 34

Table 2.2: Methods of amino acid synthesis

2.1.4 Unnatural Amino Acid Synthesis via Pd-Catalysed Negishi Cross-Coupling Synthesis

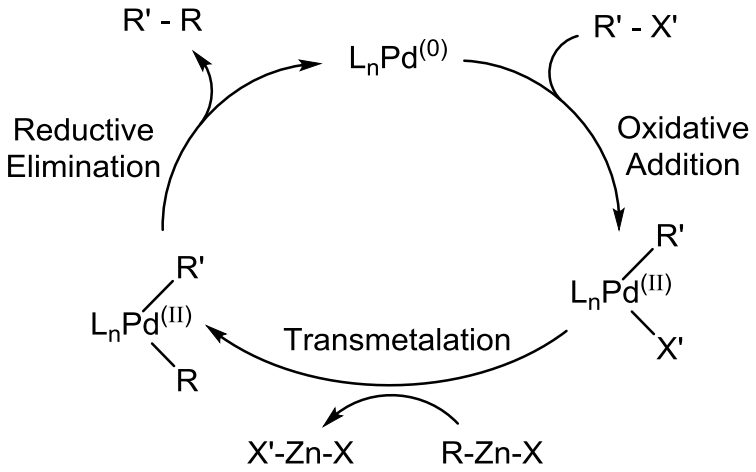
In general, Pd-catalyzed cross-coupling reactions are one of the most useful methods for C-C bond formation. Amongst them, the Negishi reaction is a mild reaction that has been shown to be well suited for the coupling of alkyl substuents.[41] Therefore, a variety of highly functional aromatic α-, β- and γ-heteroaromatic amino acids[42] have been synthesised *via* the coupling of simple iodo-amino acid building blocks and halogenated aromatics.[32] In general, the iodo-amino acids are readily synthesised in one step from natural hydroxyl-amino acids (e.g. Ser, Thr and Hyp). However, the Negishi reaction is moisture sensitive, limited to moderate functional group tolerance and the use of undesirable excess Zinc.[43] Selected examples of the amino acids that have been synthesised from L-serine using the Pd-catalysed Negishi cross-coupling approach by the Cobb group and others are shown in **Scheme 2.2**.[33,44]



Scheme 2.2: *i*: Zn, I_2, R_1-X, $Pd^{(0)}$, L_nP, DMF. A selection of heteroaromatic/aromatic amino acids containing interesting functional groups (R_1). All Compounds were synthesised from a common protected (P_1 and P_2 as *t*-butyloxycarbonyl (Boc) and benzyl (Bzl) protection, respectively) β-iodo amino acid using a Pd-catalysed Negishi cross-coupling.

The Negishi cross-coupling reaction combines organo-zinc reagents with unsaturated bromides or iodides to form a carbon–carbon bond under mild conditions using a catalytic metal from the Ni triad (Ni, Pd and Pt).[32] However, the metal of choice in Negishi chemistry is generally a $Pd^{(0)}$ species. Pt as a catalyst is not synthetically useful as the stability of the $R'RPt^{(II)}L_n$ state is high and thus, the subsequent reductive elimination step is significantly slow.[45] Ni is the least expensive member of the triad and readily used as a catalyst in Negishi cross-couplings as Ni is particularly active in the formation of alkyl-alkyl bonds.[46] However, the Ni reactions are notably less clean with lower levels of regio- and stereo-control than the Pd catalysed reactions. This is due to the ability of Ni to access one electron redox processes (unlike Pd that heavily favours +2 and 0 oxidation states)[47] which leads to various/unusual mechanisms that are not yet fully understood.[46]

The Pd-catalysed Negishi cross-coupling reaction has two major steps; the formation of the organo-zinc (R–Zn–X) intermediate from an alkyl halide and then usage of this intermediate in the subsequent sequence of oxidative addition, transmetallation and reductive elimination that form the catalytic cycle (**Scheme 2.3**).



Scheme 2.3: The sequence of oxidative addition of R'–X' to the $Pd^{(0)}$ metal centre, transmetallation with the organo-zinc intermediate and finally reductive elimination to yield the R–R' bond and regenerate L_nPd.

The Pd-catalysed Negishi cross-coupling reaction is tolerant to a broad range of functional groups; ketones, esters, nitriles, aldehydes, alkynes, amides, amines and phenols. However, phenolic[48] and amino[49] functions can significantly reduce the efficiency and yield. Acidic groups (e.g. carboxylic acids) are unfeasible as the pK_a of the acidic hydrogen is sufficient to quench of the organo-zinc reagent (R-Zn-X).[49] Therefore, in the synthesis of unnatural amino acids an appropriate orthogonal protection strategy for the amino and carboxylic acid would be beneficial. In addition, it would be advantageous if this strategy is compatible with standard Fmoc SPPS.

The pioneering efforts of the Jackson group[32] has led to the optimisation of the reaction conditions and the synthesis of large libraries of α-, β- and γ-substituted amino acids[50] from halo-iodides[51] and bromides.[52] Despite this work, the application of the Negishi reaction does still have limitations and certain compounds have proven difficult and/or not widely reported: *1*. The preparation of 5-membered heteroaromatic (furyl-/thienyl-) amino acids (**Section 2.3**). *2*. The preparation of fluorinated tyrosine amino acid derivatives (**Section 2.5**).

2.2 Bioactive 5-Membered Heteroaromatic Molecules

2.2.1 Furyl and Thienyl Moieties in Drug Molecules and Natural Products

The synthesis of 5-membered thienyl- and furyl- amino acids became a point of interest as these "unnatural" moieties are frequently found in drug molecules (**Figure 2.2**) and as ß–substituted amino acids in peptide natural products (**Figure 2.3**).

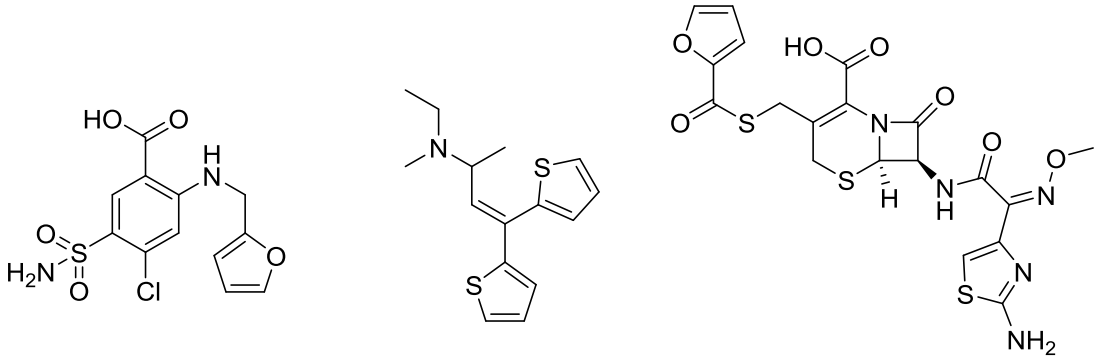


Figure 2.2: The chemical structures of the drug molecules that contain the furyl or thienyl moiety: Furosemide**,**[53,54] Ethylmethylthiambutene[55,56] and Ceftiofur.[57,58]

Furosemide is an antihypertensive agent used to treat congestive heart failure and edema,[54,59] Emethibutin is an opioid analgesic[56,60] and Ceftiofur is a third generation antibiotic used in veterinary medicine.[57,58]

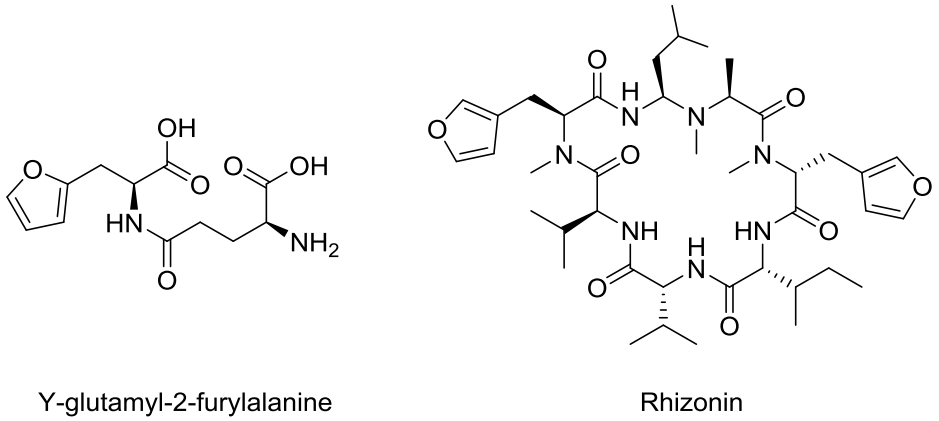


Figure 2.3: The chemical structures of the peptide natural products that contain the furyl moiety: γ-glutamyl-2-furylalanine and Rhizonin.

γ-Glutamyl-2-furylalanine[61] and rhizonin are examples of peptide natural products that contain the 2-furyl or 3-furyl moiety respectively.[62] γ-Glutamyl-2-furylalanine is produced by leaf beetle larvae as a chemical defence/feeding deterrent against insects and rhizonin is a hepatotoxic cyclo-peptide produced by the fungal strain *Rhizopus microsporus*. The bioactivity of the aforementioned non-natural aromatic amino acid containing molecules led us to investigate if these amino acids could be easily synthesised and incorporated into bioactive scaffolds. For example, as the aromatic residue in the selective CCL2 induced chemotaxis inhibiting DKPs (**Chapter 4**). To facilitate this, we attempted to access a small library of orthogonally protected thienyl- and furyl-amino acids (**Figure 2.4**)

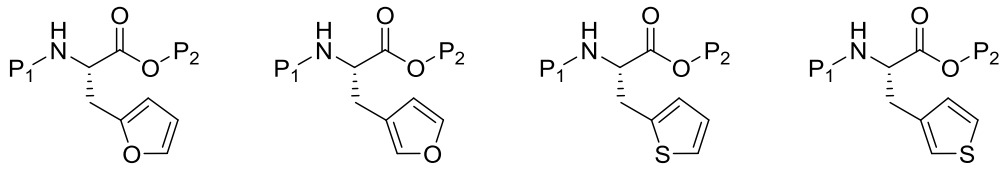
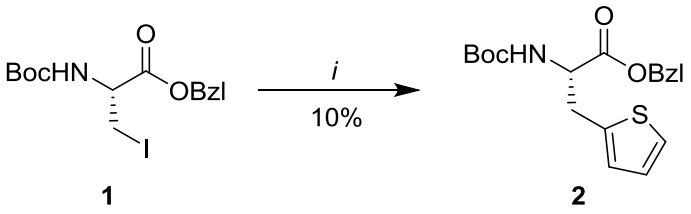


Figure 2.4: Target 5-membered β-furyl and β-thienyl amino acids

2.3 Synthesis of 5-Membered Heteroaromatic Amino Acids via Pd-Catalysed Negishi Cross-Coupling.

2.3.1 Previous Syntheses of 5-Membered Heteroaromatic Amino Acids

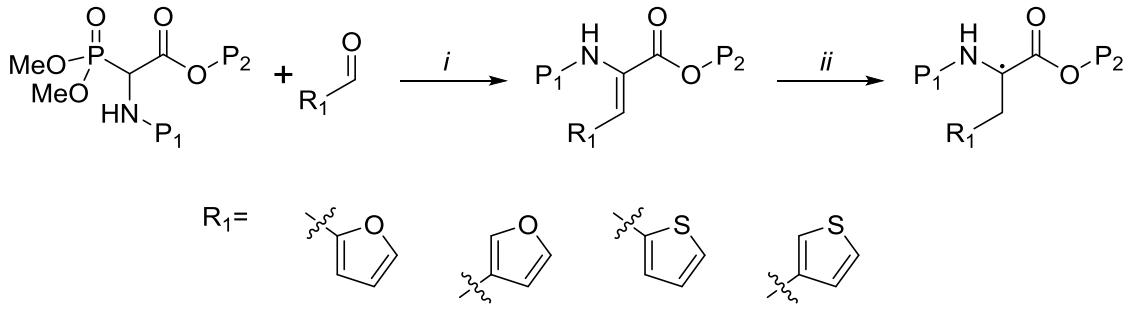
Reports of the preparation of 5-membered heteroaromatic amino acids *via* a Negishi cross-coupling are almost non-existent in the literature. Prior to this work, no published reports of the Negishi cross-coupling synthesis of furyl-amino acids exist. The only literature example was published in 1989 by the Jackson group. The report shows the synthesis of 2-thienylalanine **2** in a very poor yield (**Scheme 2.4**).[51]



1 2

Scheme 2.4: *i*: Zn/Cu couple, **1**. sonication, benzene/DMF (15:1); $[P(o\text{-}tol)_3]_2PdCl$; 2-iodothiophene, 50 °C, 1 h.[51]

Heteroaromatic amino derivatives have previously been synthesised *via* asymmetric hydrogenation (**Scheme 2.5** and Entry 1, **Table 2.2**) of dehydroamino derivative acids. Each dehydroamino acid was first synthesised *via* a Horner-Wadsworth-Emmons using furfurals and thiophene aldehydes.[63] The final asymmetric hydrogenation step yields furyl-/thienyl- alanine in excellent yield and in an enantiomeric excess of over 80% under optimised conditions.[64] Optimisation includes: H_2 pressure, temperature, catalyst makeup (metal and ligands), catalyst loading and *N*- (P_1) and *C*-(P_2) substrate protecting groups.[36,63]



Scheme 2.5: *i*: KO*t*Bu, DCM, -78° to RT. *ii*: H_2 (5 bar), [Rh], MandyPhos, MeOH, toluene (1:1)

This hydrogenation synthetic route is viable. However, the development of Negishi cross-coupling method would be beneficial as the strategy is shorter, does not use elevated H_2 pressure, requires no chiral purification and large libraries can be synthesised from a common iodo-amino acid building block (**Section 2.3.2.2**) and commercially available halo-aromatics (**Section 2.3.2.1**).

2.3.2 Preparation of Starting Materials

2.3.2.1 Halo-Aromatic

A variety of affordable commercially available 5-membered halo-iodide/bromide reagents are readily available and libraries of heteroaromatic amino acids could easily be produced. **Figure 2.5** shows the broad range of commercially available reagents (Sigma-Aldrich® and Maybridge Chemical Co.)

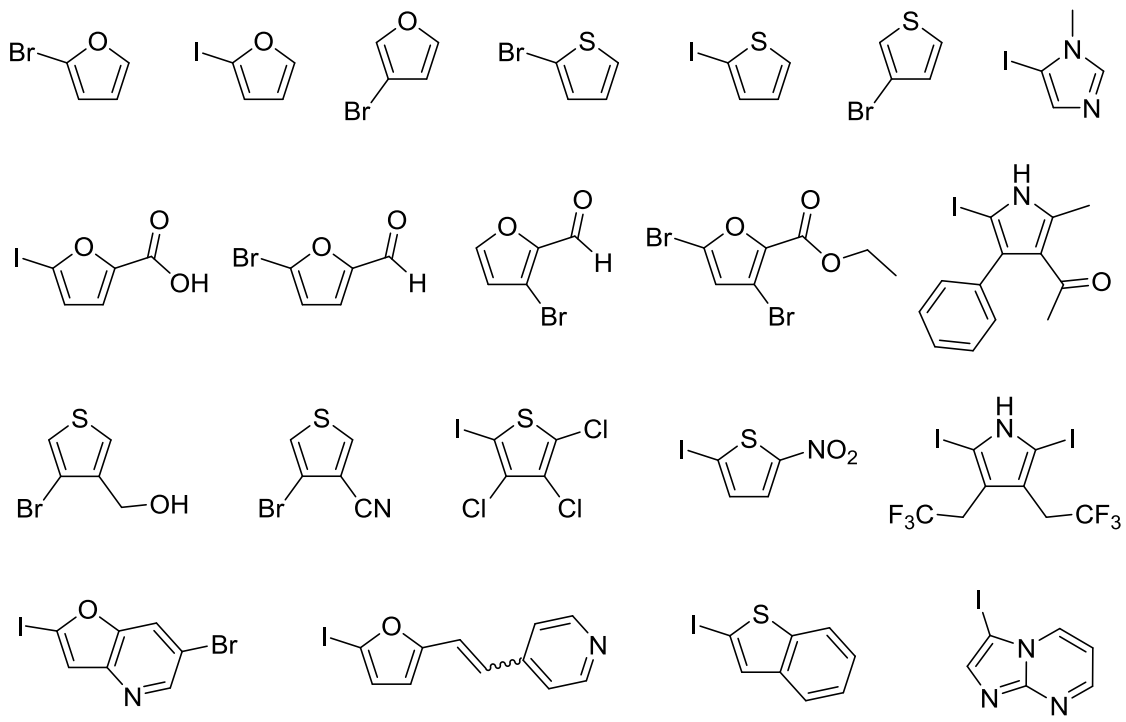
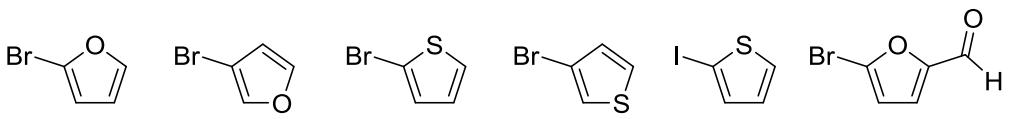


Figure 2.5: Selected examples of commercially available (Sigma-Aldrich® and Maybridge Chemical Co.) bromo or iodo substituted 5-membered heteroaromatics.

To investigate the scope of the Negishi cross-coupling reaction in relation to the synthesis of furyl- or thienyl- amino acids we selected a small set of six different halo-aromatic starting materials (**Figure 2.6**).



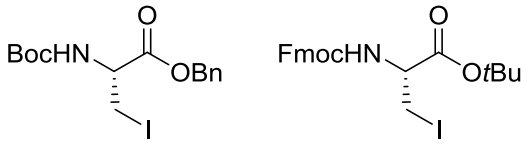
3 **4** **5** **6** **7** **8**

Figure 2.6: Selected commercially available halo-aromatic starting materials used in this study.

The 2-bromofuran (**3**), 3-bromofuran (**4**), 2-bromothiophene (**5**) and 3-bromothiophene (**6**) aromatics are a simple set of derivatives that will offer information on both how substitution position (2 versus 3) and heteroatom identity (O versus S) affects the cross-coupling. The effect of the halogen identity (Br versus I) will also be investigated by comparing the cross-coupling of 2-iodothiophene (**7**) with **6**. 2-bromofuraldehyde (**8**) will show the effect of an electron withdrawing aromatic substituent.

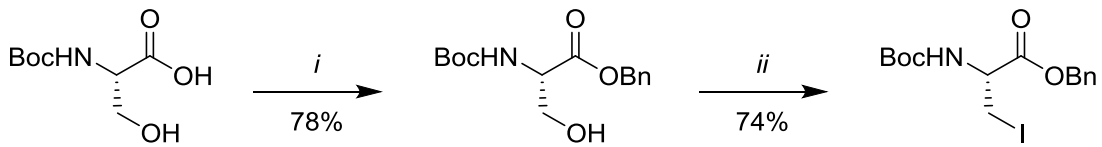
2.3.2.2 β-Iodo Amino Acid

β-Iodo amino acids are routinely and efficiently synthesised from β-hydroxyl amino acids (Ser, Thr) with various protection strategies.[65] The protection strategy of choice is guided to afford products that are usable in SPPS with minimal subsequent transformations. The two major SPPS strategies are defined by their *N*-terminal protection: Boc or Fmoc. Therefore, the iodoalanine building blocks that we synthesised combined either *N*-terminal Boc or Fmoc protection with orthogonal carboxylic acid protection (e.g. benzyl (Bzl) and *t*-butyl (*t*Bu), respectively). **Figure 2.7**. The syntheses of Boc-β-iodo-Ala-OBzl (**1**) and Fmoc-β-iodo-Ala-O*t*Bu (**9**) are described in **Scheme 2.6** and **Scheme 2.7**, respectively.



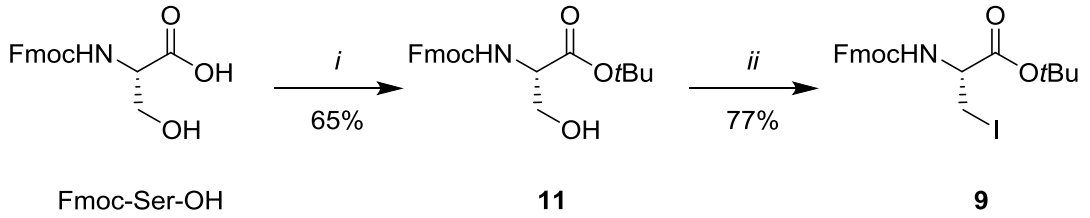
1 **9**

Figure 2.7: Boc-β-iodo-Ala-OBzl (**1**) and Fmoc-β-iodo-Ala-O*t*Bu (**9**)



Boc-Ser-OH **10** **1**

Scheme 2.6: Synthesis of **1**. *i*: Benzyl bromide, K_2CO_3, DMF, RT, 18 h. *ii*: I_2, PPh_3, imidazole, DCM, RT, 18 h.



Fmoc-Ser-OH **11** **9**

Scheme 2.7: Synthesis of **9**. *i*: TBTA, ethyl acetate, RT, 18 h. *ii*: I_2, PPh_3, imidazole, DCM, RT, 18 h.

The Negishi cross-coupling has been used to synthesise aromatic amino acids since 1989 and the vast majority of cases utilise *N*-Boc protection.[32,44,51,52,66-68] Negishi cross-coupling with *N*-Fmoc protected starting materials is a recent development in the literature and there is only a small selection of examples that exist.[69-72] Most reports show minimal differences in coupling efficiency (% yield) between *N*-Fmoc and *N*-Boc protection under the same conditions and Jackson states that "Fmoc protection is not a limiting factor in the coupling process" regarding Entry 1 in **Table 2.3**.[70] However, some of the few comparative reports show a significant lowering of the yield when *N*-Fmoc protection is used (Entries 2, 4 and 7 in **Table 2.3**).[73,74]

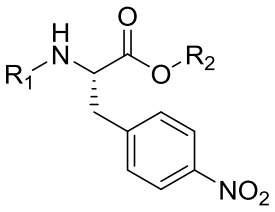
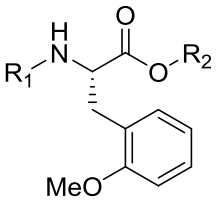
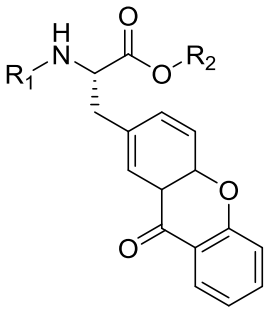
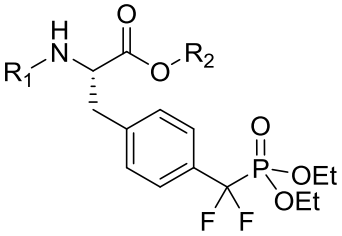
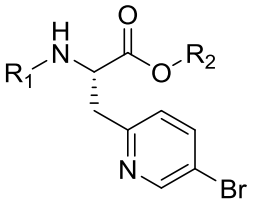
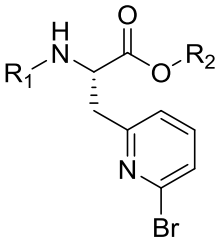
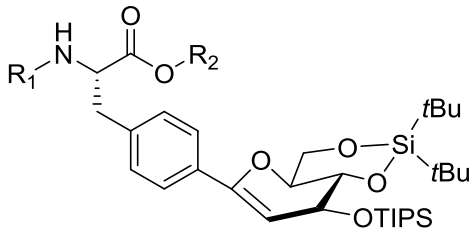
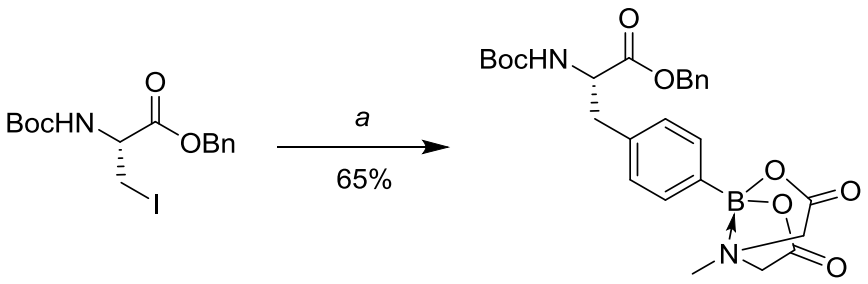
Entry	Product Structure	Sub[n]	R_1	R_2	Yield (%)	Ref
1		-	Boc	Bzl	57	66, 70
			Fmoc	*t*Bu	59	
2		-	Boc	Bzl	57	66, 70
			Fmoc	*t*Bu	31	
3		-	Boc	Me	40	75
			Fmoc	Me	42	
4		-	Boc	Me	56	73
			Fmoc	Me	32	
5		2,6	Boc	Me	52	72
			Fmoc	Me	57	
6		2,5	Boc	Me	45	
			Fmoc	Me	60	
7		-	Boc	*t*Bu	90	74
			Fmoc	Hept	69	

Table 2.3: A comparison of the effects of *N*-protection (R_1) on a Negishi cross-coupling.

It should be noted that in modern SPPS chemistry the Fmoc strategy is more prevalent. Hence, aromatic amino acids obtained from **9** are the focus of our work. However, **Section 2.3.5** describes a small comparative study of the syntheses of two alternatively protected 2-furyl heteroaromatic amino acid derivatives to investigate the effects of Boc/Bzl and Fmoc/O*t*bu amino acid protection.

2.3.3 Syntheses of Heteroaromatic N-Fmoc Amino Acids via Pd-Catalysed Negishi Cross-Coupling.

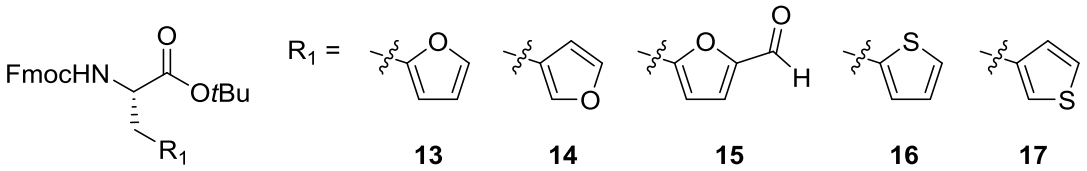
With the β-iodo amino acids and selected halo-aromatics in hand, the Negishi cross-coupling reactions were undertaken. Previous work in our group developed working conditions for the cross-coupling reaction (**Scheme 2.8**).[33]



1 **12**

Scheme 2.8: *a*: Zn (4.00 eq), 100 °C, 20 min, *in vacuo*; I_2 (cat.), DMF, 70 °C, 20 min, argon; **1**, 50 °C, 20 min; ArX (1.00 eq), $Pd_2(dba)_3$ (0.03 eq), $P(o\text{-}tol)_3$ (0.10 eq), DMF, 50 °C, 5 h; RT, 18 h.

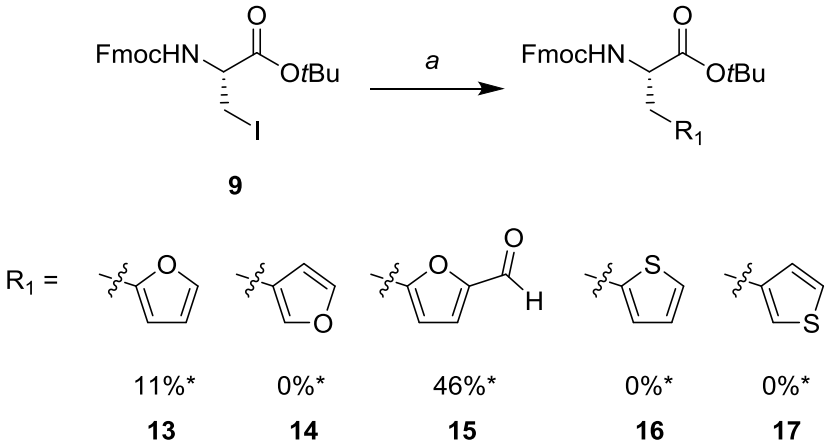
The conditions stated above (*a* in **Scheme 2.8**) became a starting point to attempt to access the desired *N*-Fmoc 5-membered furyl- and thienyl-amino acids (**Figure 2.8**).



13 **14** **15** **16** **17**

Figure 2.8: Target furyl- and thienyl- amino acids.

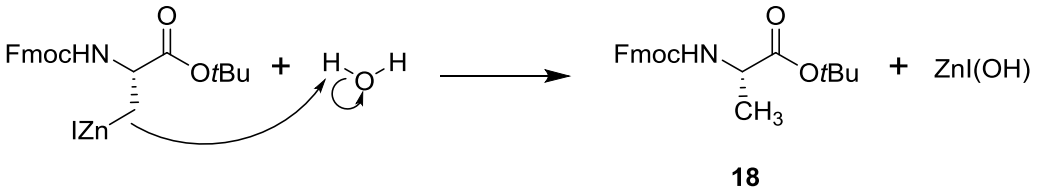
However, under reactions conditions: (*a*) (**Scheme 2.8** and **Scheme 2.9**) the only reactions that led to purified product were with **13** and **15**. Compounds **14**, **16** and **17** were visible as product peaks in crude ESI+ mass spectrum but extensive purification did not lead to isolatable product and in every case a good yield of an impurity: Fmoc-Ala-O*t*Bu (**18**) was isolated (<70%).



Scheme 2.9: *a*: Zn (4.00 eq), 100 °C, 20 min, *in vacuo*; I_2 (cat.), DMF, 70 °C, 20 min, argon; **9**, 50 °C, 20 min; ArX **3**, **4**, **5**, **6** or **8** (1.00 eq), $Pd_2(dba)_3$ (0.03 eq), $P(o\text{-}tol)_3$ (0.10 eq), DMF, 50 °C, 5 h; RT, 18 h. *Denotes isolated yield of pure product.

Previous studies have found the formation of the alanine (product of protonation of the organo-zinc intermediate) as well as a β-alkenyl amino acid (product of ß-hydride elimination) to be common reaction impurities.[66,76] In the above examples the alanine impurity was the major by-product and the β-alkenyl amino acid formation was not observed in any case.

The significant formation of the alanine impurity could stem from the slow reaction rates of the halo-furan and halo-thiophene derivatives with the Pd metal catalyst. Hence, when the uncompleted reaction is worked up it is in contact with moisture from air and silica and any residual organozinc amino acid would be quickly quenched, forming alanine: **18** in **Scheme 2.10**.



Scheme 2.10: The quenching of the organo-zinc intermediate to form **18**.

Another plausible explanation for the limited yields are that the rate of formation of **18** is on a similar or faster timescale to the product formation and thus a mixture of these compounds are synthesised during the reaction.[66] This would be the case if external proton sources entered the reaction vessel e.g. H_2O. However, in all reactions attempts were made to reduce moisture content: the reaction was carried out with anhydrous solvents and in an argon atmosphere. No experiments have been undertaken to identify the specific pathway(s) to impurity formation. However, numerous attempts have been made to optimise the reaction conditions to enable the synthesis of all target molecules (**Section 2.3.4**).

2.3.4 Optimisation of the Synthesis of Heteroaromatic N-Fmoc Amino Acids

Jackson's group showed that a significant increase in yields as well as a wider scope of starting materials could be achieved on switching phosphine ligands from P(O-tolyl)$_3$ to SPhos, the structures of which are shown below in **Figure 2.9**.[32] SPhos is a very bulky biaryl ligand that is known to improve various Pd cross-coupling reactions, as seen with the Suzuki-Miyaura reaction.[77,78] SPhos enabled the synthesis of extremely hindered aryl boronic acids and aryl halides and also allowed for reduced catalyst loading.[32,77]

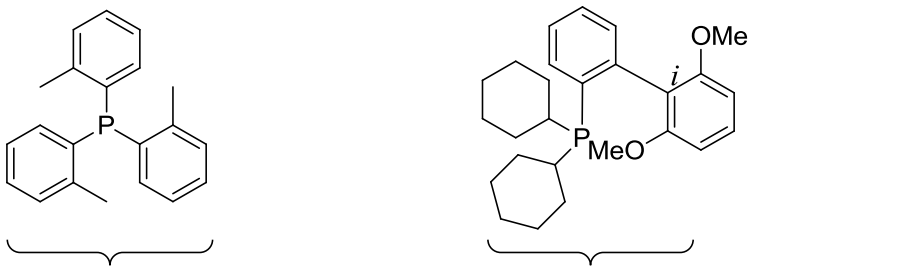
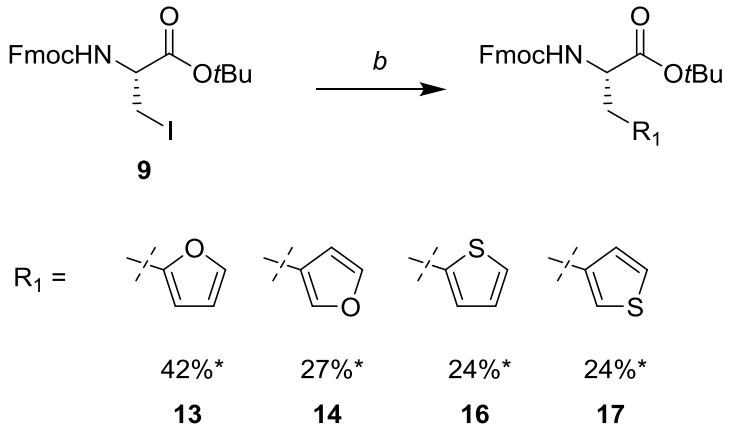


Figure 2.9: The chemical structures of P(O-tolyl)$_3$ to SPhos phosphine ligands.

Literature suggests that in the cross-coupling catalytic cycle (**Scheme 2.3**) the monoligated phosphine palladium species is the most reactive.[79] Biarylphosphines like SPhos have the ability to stabilise both $Pd^{(0)}$ and $Pd^{(II)}$ monoligated species. The stabilisation of the $Pd^{(0)}$ species is due to interactions between the $Pd^{(0)}$ centre and the π-aromatic system (specifically the *ipso*-carbon (*i* in **Figure 2.9**). The stabilisation of the $Pd^{(II)}$ species is *via* coordination of the alkoxy (OMe) substituent.

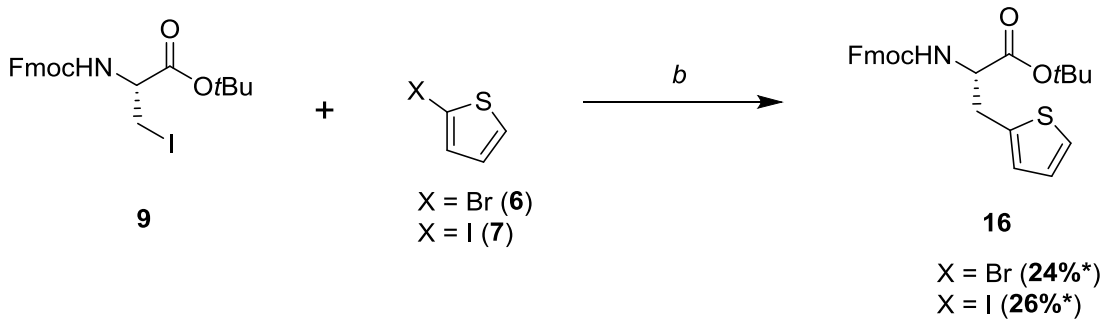
The rate limiting step in the analogous Suzuki-Miyaura reaction is believed to be the transmetallation step.[80] Therefore, a huge benefit is gained by the stabilisation of the product from the oxidative addition step (The monoligated $Pd^{(II)}$ species) as this subsequently performs the "rate-limiting" transmetallation.[78] SPhos is a bulky ligand with excellent properties for Pd cross-coupling chemistry.[32] Hence, the synthesis of low yielding compounds **13, 14, 16** and **17** was undertaken using reaction conditions (*b*) that made use of the SPhos ligand (**Scheme 2.11**).



Scheme 2.11: *b*: Zn (4.00 eq), 100 °C, 20 min, *in vacuo*; I_2 (cat.), DMF, 70 °C, 20 min, argon; **9**, 50 °C, 20 min; ArX **3**, **4**, **5** or **6** (1.00 eq), $Pd_2(dba)_3$ (0.03 eq), SPhos (0.09 eq), DMF, 50 °C, 5 h; RT, 18 h. *Denotes isolated yield of pure product.

Reaction conditions (*b*) showed significant improvement over conditions (*a*), allowed access to compounds **14**, **16** and **17** and improved the yield of **13**. However, the yields are still only moderate (24 - 42%). Therefore, further efforts were made to increase yield: the use of aromatic iodides and temperature lowering.

In general, aromatic iodides are found to be better reagents for cross-coupling reactions than the corresponding aromatic bromides.[52] An increase in yield when using iodo-aromatics is generally expected as they typically have lower bond dissociation energies than bromo-aromatics. This leads to an increased reactivity in the oxidative addition step of the catalytic cycle.[52] The severity of the effect can vary significantly and has shown to be more pronounced with sterically hindered or phenolic halo-aromatics.[52] Therefore, to further improve yields we attempted a trial reaction with commercially available 2-iodothiophene (**7**), under conditions; *b* in **Scheme 2.12**.

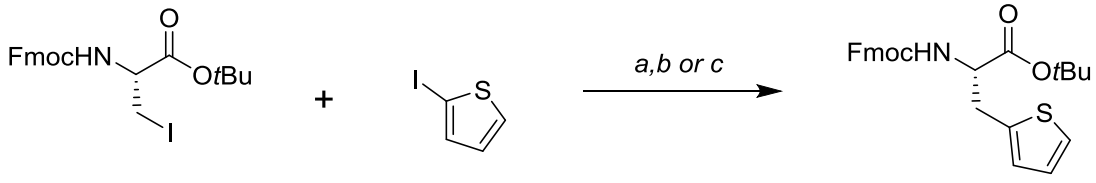


Scheme 2.12: *b*: Zn (4.00 eq), 100 °C, 20 min, *in vacuo*; I_2 (cat.), DMF, 70 °C, 20 min, argon; **9**, 50 °C, 20 min; ArX **6** or **7** (1.00 eq), $Pd_2(dba)_3$ (0.03 eq), SPhos (0.09 eq), DMF, 50 °C, 5 h; RT, 18 h. *Denotes isolated yield of pure product.

As indicated in **Scheme 2.12** the observed yield increase when using iodothiophene was small (24% to 26%) and due to the restricted set of commercially available iodo-furans further investigations were not carried out.

The final attempt to optimise the cross-coupling was to conduct a trial reaction at lower temperature as the temperatures used in conditions *a* and *b* could possibly contribute to a lowering of yields *via* compound breakdown. However, previous studies found a balance between increasing reaction rates and thermal decomposition was needed. Generally, reactions conducted at 50 °C yielded optimal results.[52,66]

This trial was undertaken using 2-iodothiophene with the reaction temperature for the first 5 hours of the reaction lowered from 50 °C to room temperature: conditions *c* (**Scheme 2.13**).



9 **16**

a = **0%***
b = **26%***
c = **22%***

Scheme 2.13:Synthesis of **16** under conditions *a*, *b* and *c*. *a*: Zn (4.00 eq), 100 °C, 20 min, *in vacuo*; I_2 (cat.), DMF, 70 °C, 20 min, argon; **9**, 50 °C, 20 min; $Pd_2(dba)_3$ (0.03 eq), $P(o\text{-}tol)_3$ (0.10 eq), **7** (1.00 eq), DMF, 50 °C, 5 h; RT, 18 h. *b*: Zn (4.00 eq), 100 °C, 20 min, *in vacuo*; I_2 (cat.), DMF, 70 °C, 20 min, argon; **9**, 50 °C, 20 min; $Pd_2(dba)_3$ (0.03 eq), SPhos (0.09 eq), **7** (1.00 eq), DMF, 50 °C, 5 h; RT, 18 h. *c)* Zn dust (4.00 eq), 100 °C, 20min, *in vacuo*; I_2 (cat.), DMF, 50 °C, 20min, argon; **9**, RT, 20 min; $Pd_2(dba)_3$ (0.03 eq), SPhos (0.10 eq), **7** (1.00 eq), argon, DMF, RT, 24 h. *Denotes isolated yield of pure product.

The trail reaction at lower temperature gave rise to a slightly lower yield (22%). This agrees with previous findings.[52] This is evidence that thermal decomposition during the reaction is not a significant yield limiting factor. **Table 2.4** summarises the entire work in this area.

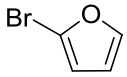
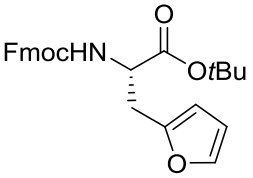
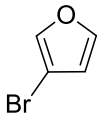
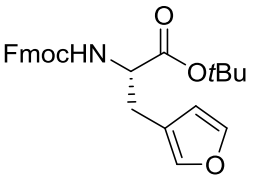
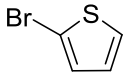
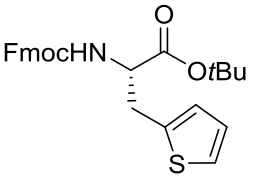
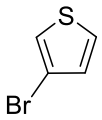
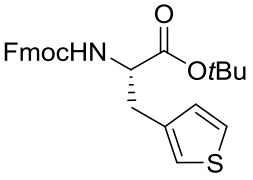
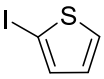
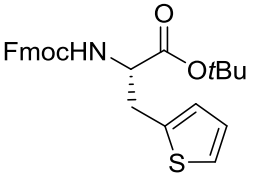
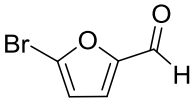
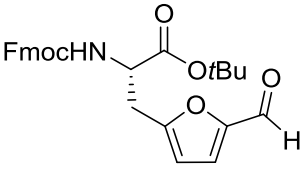
Entry	Halo-aromatic (ArX)	Product	Conditions Yield (%)* a:	b:	c:
1	 **3**	 **13**	11	42	-
2	 **4**	 **14**	0	27	-
3	 **5**	 **16**	0	24	-
4	 **6**	 **17**	0	24	-
5	 **7**	 **16**	0	26	22
6	 **8**	 **15**	46	-	-

Table 2.4: Summary of Negishi cross-coupling reactions to form thienyl- or furyl-alanines under conditions: *a*, *b* and *c*. *a*: Zn (4.00 eq), 100 °C, 20 min, *in vacuo*; I_2 (cat.), DMF, 70 °C, 20 min, argon; **9**, 50 °C, 20 min; $Pd_2(dba)_3$ (0.03 eq), P(o-tol)$_3$ (0.10 eq), ArX (1.00 eq), DMF, 50 °C, 5 h; RT, 18 h. *b*: Zn (4.00 eq), 100 °C, 20 min, *in vacuo*; I_2 (cat.), DMF, 70 °C, 20 min, argon; **9**, 50 °C, 20 min; $Pd_2(dba)_3$ (0.03 eq), SPhos (0.09 eq), ArX (1.00 eq), DMF, 50 °C, 5 h; RT, 18 h. *c:* Zn dust (4.00 eq), 100 °C, 20min, *in vacuo*; I_2 (cat.), DMF, 50 °C, 20min, argon; **9**, RT, 20 min; $Pd_2(dba)_3$ (0.03 eq), SPhos (0.10 eq), ArX (1.00 eq), argon, DMF, RT, 24 h. *Denotes isolated yield of pure product.

In the Pd-catalysed cross-coupling reaction to form compounds **13** – **17** in **Table 2.4** a major difference in the isolatable yields was observed by switching phosphine ligands. However, the yield was also dependant on the halo-aromatic and the substitution positions. The synthesis of the furan (**13** and **14**) and thiophene (**16** and **17**) amino acid derivatives gave yields ranging from 22 to 46%. This is potentially due to the electronic properties of the aromatic rings and can be visualised by their NMR spectra **Figure 2.10**.

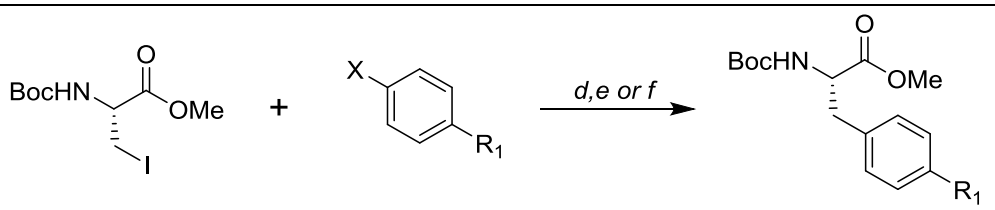
furan **thiophene**

Figure 2.10: The chemical shift in ppm (in $CDCl_3$) of the heteroaromatic functional group in furan and thiophene. In both cases the electron density is lower at the 2 position (less deshielded).

Hydrogen atoms at positions 2, 5 have decreased electron density over the 3, 4. The electron withdrawing inductive effect of the significantly more electronegative oxygen or slightly more electronegative sulphur atom is pronounced at the 2 position. The electron deficiency at the 2 position (**Figure 2.10**) translates to the increased yields compared to the 3 position. The yield of **13** (42%) is increased compared to **14** (27%) under the same conditions. The effect is more pronounced for the furan than the thiophene derivatives because the oxygen atom is more electronegative than sulphur. These results agree with the literature surrounding the Suzuki-Miyaura reaction as evidence suggests that the effect of an electron withdrawing groups (EWG) is to increase the reactivity of Pd-catalysed cross-coupling due to an enhancement of the oxidative addition step.[78] Therefore, an improvement in the cross-coupling reaction of halo-aromatics would be observed at positions on the ring with lower electron density (e.g. position 2) or when EWGs are present (e.g. an aldehyde). **15** was synthesised in moderate yield (46%) without using the SPhos ligand as the aldehyde EWG increases the reactivity of the halo-heteroaromatic (Entry 3, **Table 2.4**).[81]

However, in the synthesis of 6-membered aromatic amino acids *via* the Pd-catalysed Negishi cross-coupling an electrostatic trend based on the halo-aromatic is not

observed. The literature shows only that electron withdrawing (R_1 = F) and electron donating (R_1 = OMe) halo-aromatic substituents lower the yield compared to halo-benzene (**Table 2.5**).[82]

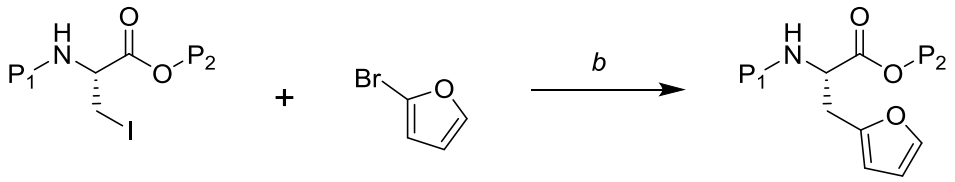


Entry	X	R_1	Conditions Yield (%)* d:[32]	e:[52]	f:[32]
1	I	H	80		
2	I	F	77		
3	I	OMe	65		
4	Br	H		60	
5	Br	F		40	
6	Br	OMe		48	
7	Br	H			77
8	Br	F			69
9	Br	OMe			65

Table 2.5: The effect of p-fluoro (EWG) and p-methoxy (EDG) groups on the Pd-catalysed Negishi cross-coupling of 6-membered halo-aromatics. (adapted from Ross *et al.*[32] and Oswald *et al.*[52]). Conditions: *d*: $Pd_2(dba)_3$ (0.25 mol %), SPhos (0.5 mol %), DMF, RT, overnight.[32] *e*: $Pd(OAc)_2$ (5.0 mol %), $P(o\text{-}tol)_3$ (10 mol %), DMF, 50 °C, 2 h.[52] *f*: $Pd_2(dba)_3$ (2.5 mol %), SPhos (5.0 mol %), DMF, 50 °C, 3 h.[32]

2.3.5 Synthesis of Heteroaromatic N-Boc Protected 2-Furyl Alanine

Given the low yields observed in the synthesis of *N*-Fmoc amino acids, we decided to synthesise the *N*-Boc-2-furylalanine-OBzl (**19**) building block as a small comparative study to probe the effect of changing the protecting groups on the Negishi cross-coupling reaction. The synthetic conditions (*b* in **Scheme 2.13** and **Scheme 2.14**) were applied to an *N*-Boc starting material (**1**) and 2-bromofuran (**3**).



1: P_1=Boc, P_2=Bzl
9: P_1=Fmoc, P_2=*t*Bu

3

19: P_1=Boc, P_2=Bzl (**86%***)
13: P_1=Fmoc, P_2=*t*Bu (**42%***)

Scheme 2.14: Synthesis of the *N*-Boc and *N*-Fmoc furylalanines **19** and **13**. *b*: Zn (4.00 eq), 100 °C, 20 min, *in vacuo*; I_2 (cat.), DMF, 70 °C, 20 min, argon; **1** or **9**, 50 °C, 20 min; **3** (1.00 eq), $Pd_2(dba)_3$ (0.03 eq), SPhos (0.09 eq), DMF, 50 °C, 5 h; RT, 18 h. *Denotes isolated yield of pure product.

This synthesis outlined in **Scheme 2.14** shows that Boc/OBzl protected amino acids are readily synthesised and show significantly higher yields when compared to the Fmoc/O*t*Bu derivative. The *N*-Boc protected product (**19**) was then used in an attempt to synthesise γ-glutamyl-2-furylalanine.

2.3.6 Synthesis of γ-Glutamyl-2-furylalanine

γ-Glutamyl-2-furylalanine is a simple dipeptide natural product containing L-(2-furyl)alanine (**Figure 2.11**).

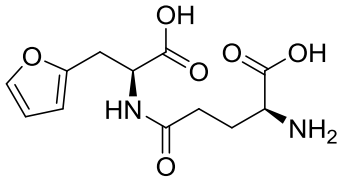
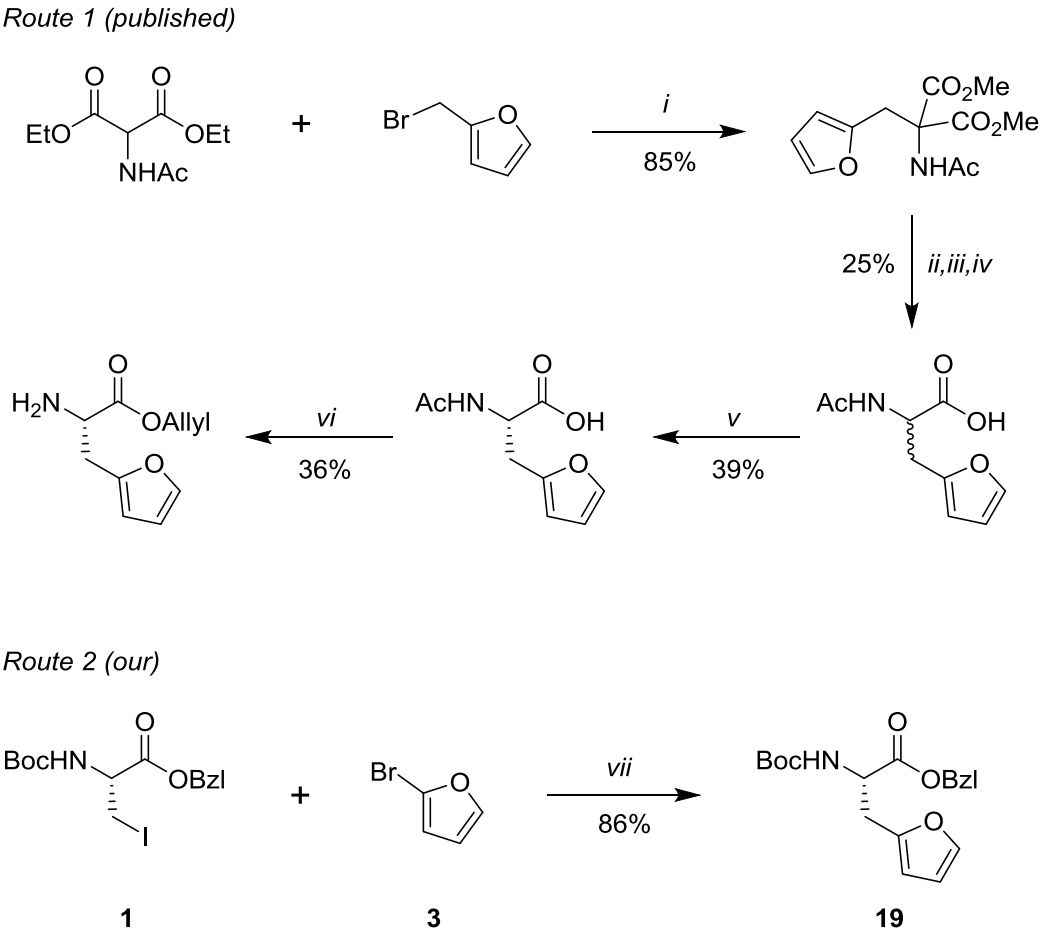


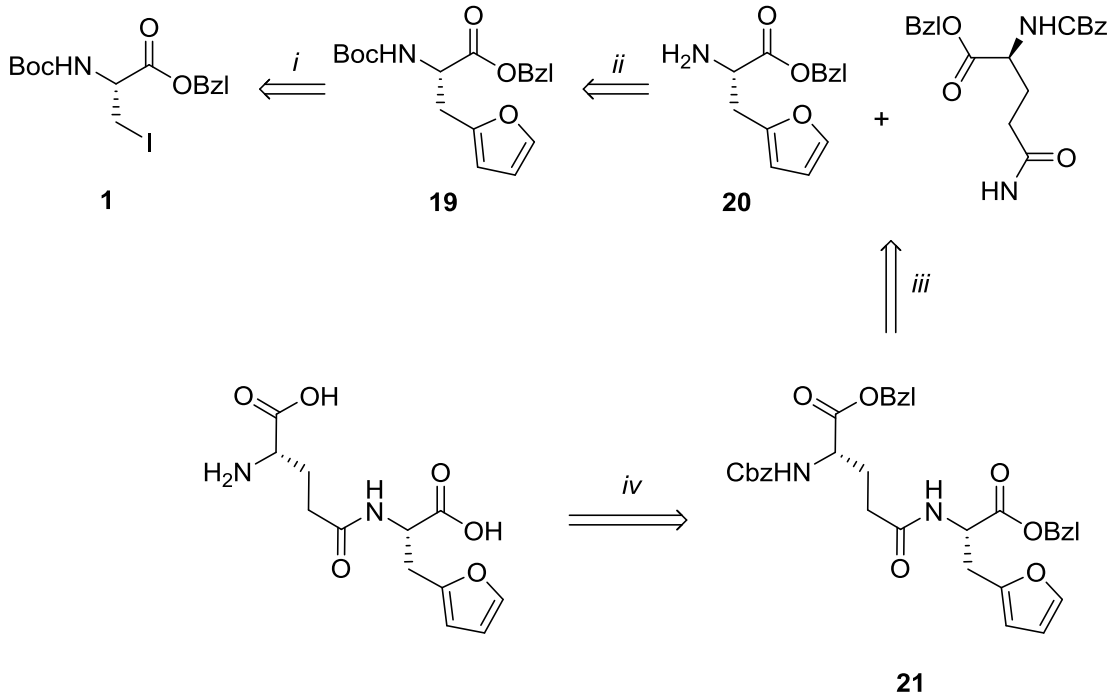
Figure 2.11: Dipeptide natural product: γ-glutamyl-2-furylalanine.

The published 10 step synthesis of γ-glutamyl-2-furylalanine is surprisingly long with 6 steps devoted to the synthesis of appropriately protected L-(2-furyl)alanine *via* an alkylation, saponification, decarboxylation and subsequent chiral resolution (*Route 1* in **Scheme 2.15**).[61] The Pd-catalysed Negishi cross-coupling reaction gives access to L-2-furylalanines commercially available iodoalanine derivatives in one step (*Route 2* in **Scheme 2.15**).



Scheme 2.15: Two routes to protected 2-furyl amino acid derivatives. *i*: NaOMe, reflux, 2 h. *ii*: $Ba(OH)_2$, RT, 24 h. *iii*: dioxane, reflux, 24 h. *iv*: KOH, RT, 17 h. *v*: acylase I, $CoCl_2$. *vi*: allyl alcohol, *p*-TSA, benzene, reflux, 18 h. *vii*: Zn, 100 °C, 20 min, *in vacuo*; I_2, DMF, 70 °C, 20 min, argon; **1**, 50 °C, 20 min; **3**, $Pd_2(dba)_3$, SPhos, DMF, 50 °C, 5 h; RT, 18 h.

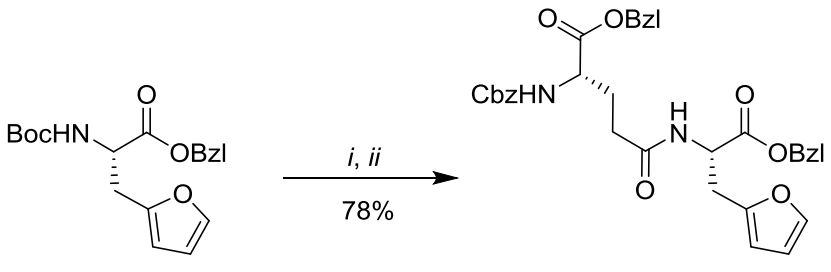
Therefore, a truncated synthesis of γ-glutamyl-2-furylalanine could be envisaged. We attempted a four step retrosynthetic pathway using commercially available building blocks (**Scheme 2.16**).



1 **19** **20** **21**

Scheme 2.16: Proposed four step retrosynthesis of γ-glutamyl-2-furylalanine. *i*: Zn (4.00 eq), 100 °C, 20 min, *in vacuo*; I_2 (cat.), DMF, 70 °C, 20 min, argon; **1**, 50 °C, 20 min; **3** (1.00 eq), $Pd_2(dba)_3$ (0.03 eq), SPhos (0.09 eq), DMF, 50 °C, 5 h; RT, 18 h. *ii*: TFA/DCM (1:1), RT, 30 min. *iii*: Z-Glu-OBzl (1.0 eq), PyBOP (1.0 eq), Z-Glu-OBzl (1.0 eq), DIPEA (2.0 eq), DCM, RT, 20 h. *iv*: H_2, Pd(cat), MeOH, RT, 2 h.

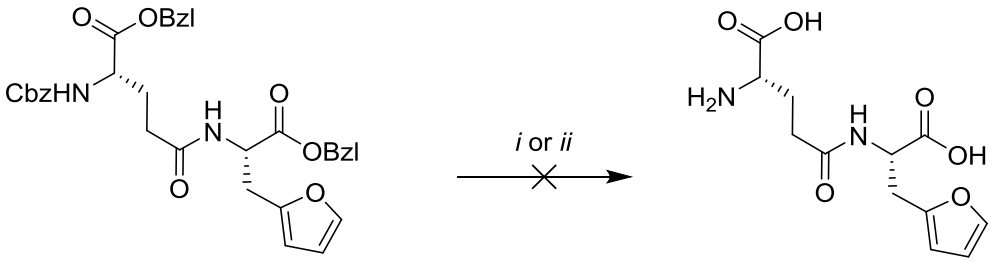
This synthetic strategy was undertaken with the first step as the acid catalysed *N*-Boc deprotection of **19**. The stability of these amino acids in synthetic procedures has been reported to be poor under acidic[83] and hydrogenation[61] conditions. In our hands, the furan moiety present proved stable to the acidic conditions and the desired dipeptide (**21**) was synthesised in a solution phase peptide coupling reaction (**Scheme 2.17**).



19 **21**

Scheme 2.17: *i*: TFA/DCM (1:1), RT, 30 min. *ii*: PyBOP (1.0 eq), Z-Glu-OBzl (1.0 eq), DIPEA (2.0 eq), DCM, RT, 20 h.

Various attempts were made to carry out the global deprotection of **21** to give natural product γ-glutamyl-2-furylalanine. However, the hydrogenation yielded a complex mixture of inseparable products with γ-glutamyl-2-furylalanine being identified only as a potential product in analysis of the crude reaction mixture by mass spectrometry (**Scheme 2.18**).



21 Y-glutamyl-2-furylalanine

Scheme 2.18: *i*: H_2, Pd(cat), MeOH, RT, 2 h. *ii*: H_2, Pd(cat), EtOAc, RT, 2 h.

This reaction shows that the benzyl protection strategy (prevalent in Negishi cross-coupling amino acid synthesis) is of limited use with furyl moieties due to their instability to standard deprotection conditions. This emphasises the importance of the work in **Section 2.3.3** to provide facile syntheses *N*-Fmoc heteroaromatic derivative amino acids as both furyl and thienyl amino acids are known to be stable to the SPPS protocols under optimised conditions.[83] Therefore, it could be envisaged the *N*-Fmoc derivatives could be used to access the natural product *via* SPPS. Alternatively, future solution phase syntheses would use iodoalanine and glutamic acid starting materials with a furyl stable global deprotection step. e.g. Boc/*t*Bu or allyl ethers.[61]

2.3.7 Heteroaromatic Amino Acids Conclusions

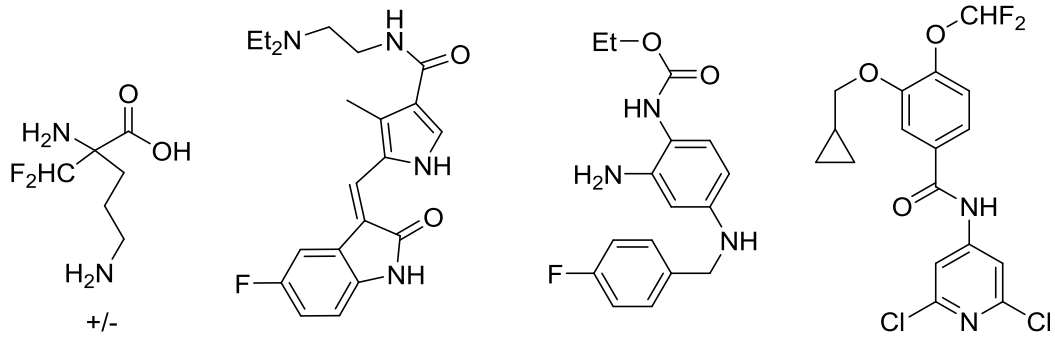
A small library of heteroaromatic amino acids that were previously difficult to access, have been prepared, **Figure 2.8**.[84] The original reaction conditions (*a* in **Scheme 2.13**) were unsatisfactory and yielded poor results. The SPhos ligand provided a route to their synthesis and gave isolatable yields of orthogonally protected heteroaromatic amino acids of 24 - 42% see **Table 2.4**. These compounds are ready to be used as building blocks in solution and SPPS e.g. on-resin synthesis of cyclic dipeptides (see **Chapter 4**).

2.4 Biologically Relevant Fluorine Containing Molecules

Fluorine is an element with a high electronegativity, small atom radius, low (C-F bond) polarizability and the natural isotope (^{19}F, 100% abundance) has a nuclear spin of ½ as well as a large magnetogyric ratio (83% that of ^{1}H).[85] These chemical and physical properties mean that the incorporation of fluorine into bioactive molecules is important in many areas of biological chemistry. Of relevance to the work reported here, are the roles that fluorine plays in medicinal chemistry (**Section 2.4.1**) and peptide/protein chemistry (**Section 2.4.2**).

2.4.1 Organofluorine in Medicinal Chemistry

Often when a C-O or C-H is substituted for a C-F bond in a bioactive molecule an increase in drug effectiveness can be observed,[86] e.g. increased metabolic stability, target binding and lipophilicity/membrane permeability.[87,88] Over the past twenty years fluorine has come to the forefront of medicinal chemistry and is second only to nitrogen as the favoured heteroatom in drug design,[87] with 9 out of the 19 US FDA approved drugs contained fluorine in 2007.[86] **Figure 2.12** shows three FDA approved drugs including: Ornidyl , a fluoro-amino acid licensed to treat sleeping sickness.[87,89]



Ornidyl Sunitinib Potiga Roflumilast

Figure 2.12: Fluorine containing drugs: Ornidyl for sleeping sickness, Sunitinib is licenced to treat gastrointestinal stromal tumor and renal cell carcinoma, Potiga for epilepsy and Roflumilast as an anti-inflammatory agent.[87]

2.4.2 Fluorine in Peptide/Protein Chemistry

Organofluorine molecules can be utilised in peptide/protein chemistry as structural probes providing the required handle to carry out ^{19}F NMR studies (**Figure 2.13**).[9,11]

Fluorine can be site selectively incorporated into bioactive molecules and the ^{19}F nuclei provides dynamic or topological NMR information in proteins (or protein complexes) which are unstable or too large for full structural analysis. ^{19}F NMR spectroscopy is particularly useful in *in vivo* measurements as there is an absence of background ^{19}F signals and the chemical shift sensitivity to the nuclei's environment is large (around a 100 fold increase in chemical shift dispersion over ^{1}H).[85] This translates to easier measurements of weak binding, folding, conformational change or reaction kinetics.[11] Therefore, ^{19}F NMR can provide insights into folded and partly folded states of proteins which is vital to understanding protein-protein interactions[90] including the mechanism of folding.[91,92]

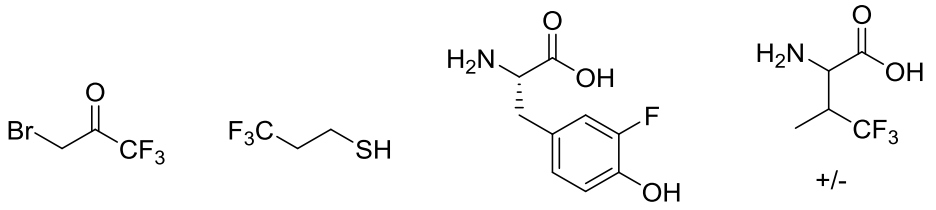
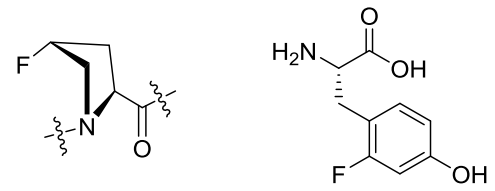


Figure 2.13: 3-Bromo-1,1,1-trifluoroacetone (BTFA) is commonly used to site selectively modify proteins in a single step reaction, e.g. in the structural elucidation of the membrane protein diacyl glycerol kinase. 2,2,2-trifluoroethanethiol (TFET) reacts specifically with cysteine thiols and has been used to measure NOEs between in the cytoplasmic domain of rhodopsin.[85] 3-Fluorotyrosine and 4,4,4-trifluorovaline are examples of ß-substituted amino acids that can be biosynthetically incorporated into bioactive molecules.[93]

Organofluorines (**Figure 2.14**) are also incorporated into to peptides/protein molecules as conformational modulators[5]/stabilizers[94] to stabilise or impart secondary structural motifs. These motifs are vital for peptide function and stability.



(2*S*, 4*R*) 4-fluoroproline (Flp)

2-fluoro-L-tyrosine

Figure 2.14: (2*S*,4*R*) 4-Fluoroproline is able to significantly strengthen a proline helix due to the fluorine gauche effect giving a high preference to an exo-position.[94] 2-Fluorotyrosine (as well as other fluorinated tyrosine derivatives) have been utilised as a way to modulate the hydrogen bonding potential and pK_a of the phenolic moiety in tyrosine.[5,18,19]

Due to the aforementioned applications, a number of fluorinated amino acids have become commercially available from multiple companies e.g. Sigma-Aldrich®, Polypeptide and High Force Research. **Figure 2.15** shows all the fluorinated amino acids currently available from Sigma-Aldrich®. The synthesis of novel fluorinated amino acids or improved synthetic routes to their synthesis is an expanding area of research.

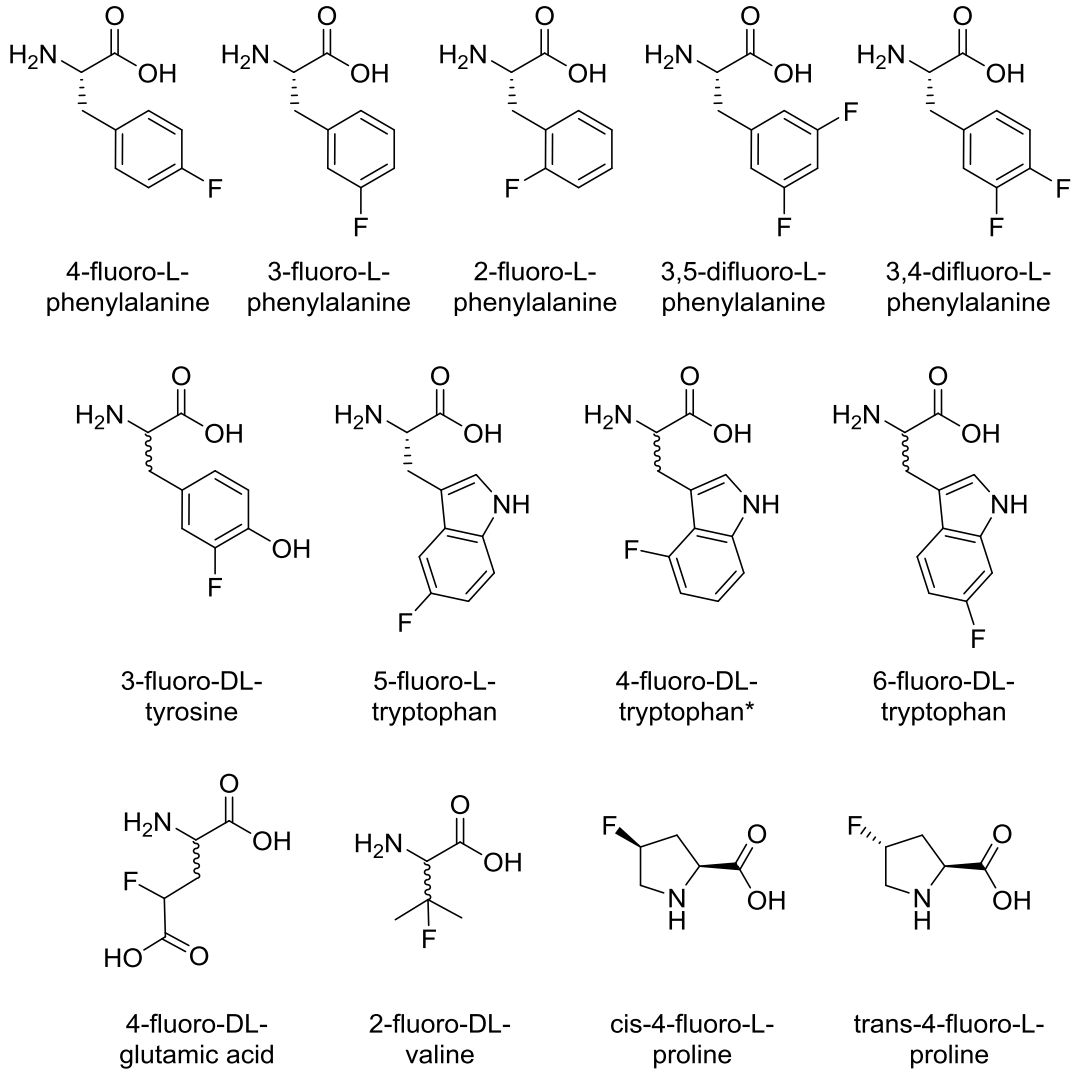


Figure 2.15: Commercially available amino acids from Sigma-Aldrich®

The Cobb group has previously prepared a range of fluorinated-amino acids with a variety of different applications: A dual IR and NMR spectroscopy probe (unpublished),[14,95] structural stabilizers[8] and in the facile synthesis of dehydroamino acid derivatives[82] (**Figure 2.16**).

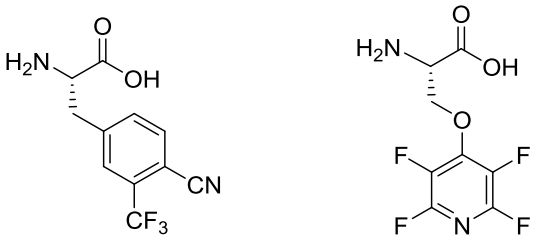


Figure 2.16: Previously synthesised fluorinated amino acids by the Cobb group: 3-trifluoromethyl, 4-cyano–L–phenylalanine and tetrafluoropyridoxy–L–serine.

2.4.3 Iturin A and Lipopeptides from Bacillus sp. CS93

Research in the Murphy group[96] at University College Dublin showed that commercially available 3-fluorotyrosine could be incorporated into the antimicrobial cyclic lipopeptide: Iturin A. Iturin A is non-ribosomally biosynthesised by *Bacillus* sp. CS93 (**Figure 2.17**).

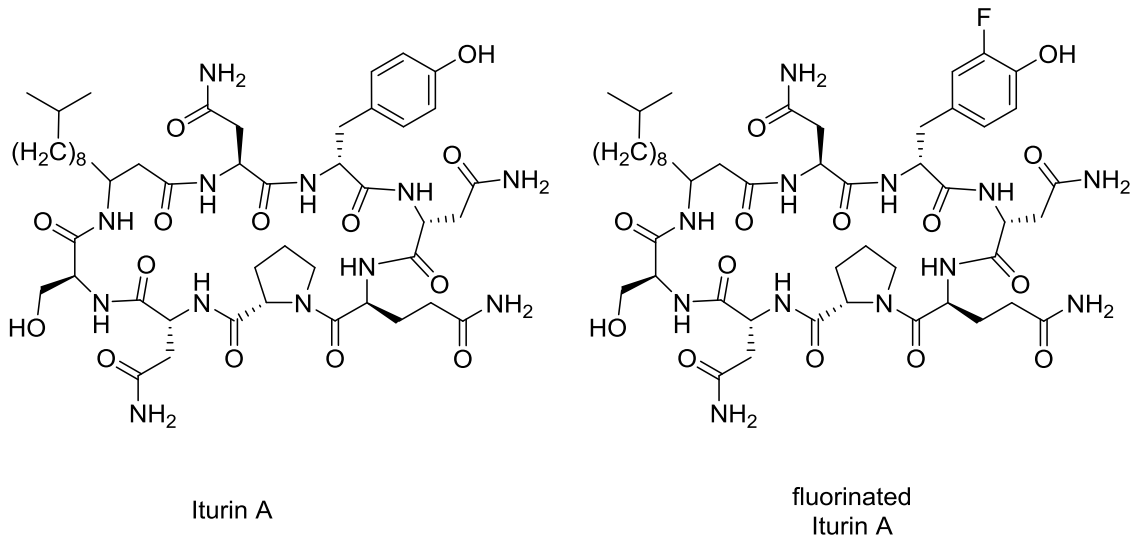


Figure 2.17: The structure of Iturin A and fluorinated Iturin A.

The antifungal mode of action of this class of lipopeptide is to generally bind and increase the permeability of the cell membrane. The hydroxyl group of the tyrosine residue of Iturin A is thought to be responsible for an important interaction in membrane binding as it interacts with fungal membrane sterol.[97] Thus, the addition of

an electron withdrawing ring substituent would alter the pK_a of the hydroxyl and affect activity (**Figure 2.18**).

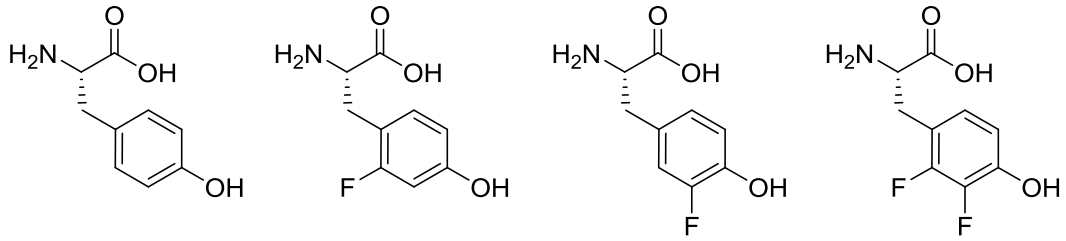


Figure 2.18: The structures of: Tyrosine, 2-fluorotyrosine (**22**), 3-fluorotyrosine and 2,3-difluorotyrosine (**23**) with the pK_a of the phenol hydroxyl indicated. pK_a were measured using spectrophotometry in aqueous buffer.[56]

In collaboration with the Murphy group we were keen to explore the potential application of the non-commercially available unprotected amino acids 2-fluorotyrosine and 2,3-difluorotyrosine (**31** and **32** in **Figure 2.18**). The aim was to use biosynthetic incorporation of these amino acids to access a small library of fluorinated lipopeptides for biological evolution.

2.5 Synthesis of Fluoro-Tyrosine Derivative Amino Acids via Pd-Catalysed Negishi Cross-Coupling.

2.5.1 Previous Syntheses of Fluorinated Tyrosine Derivative Amino Acids

The most prevalent literature route to the synthesis of fluorinated tyrosine derivatives is *via* enzymatic synthesis that utilises tyrosine phenol lyase and fluorinated phenol starting materials. Yields of purified (unprotected) amino acids range from low to moderate (10 - 42%).[19,98] Other routes do exist, e.g. *via* a chiral glycine anion equivalent.[99] Surprisingly, a Negishi cross-coupling method has not been utilised for the synthesis of fluorinated tyrosine amino acids despite the fact that a range of phenylalanine derivatives have been reported.

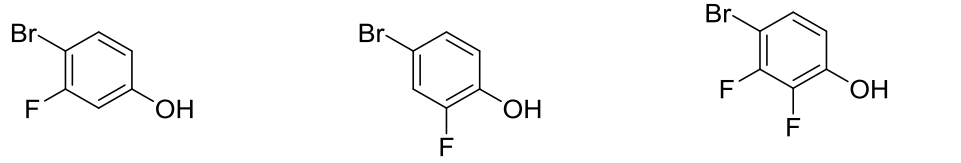
2.5.2 Preparation/Choice of Starting materials

2.5.2.1 Synthesis of β-Iodo Amino Acid

Iodo-amino acid building blocks (**1** and **9**) were previously synthesised (**Section 2.3.3** and **2.3.5**, respectively). In this study the *N*-Boc derivative (**1**) was chosen over the *N*-Fmoc (**9**) as it was shown to give higher yields and the deprotection steps yield volatile by-products that do not require an additional purification. The *N*-Boc protection is removed under acidic conditions and the OBzl is generally removed by dehydrogenation.

2.5.2.1 Preparation of Halo-Aromatic

A number of *para*-substituted halo-phenols are commercially available from Sigma-Aldrich®: Specifically, fluorinated derivatives: 4-bromo-3-fluorophenol (**24**); 4-bromo-2-fluorophenol; and 4-bromo-2,3-difluorophenol (**25**) (**Figure 2.19**).



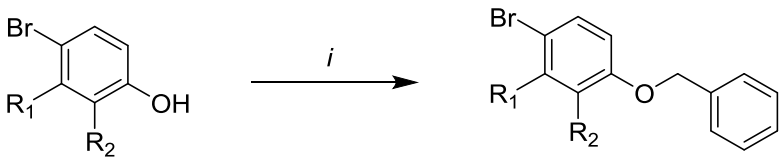
4-bromo,3-fluorophenol **24** | 4-bromo,2-fluorophenol | 4-bromo,2,3-difluorophenol **25**

Figure 2.19: The structure of: 4-bromo-3-fluorophenol (**24**); 4-bromo,2-fluorophenol; and 4-Bromo-2,3-difluorophenol (**25**)

4-Bromo-3-fluorophenol (**24**) and 4-bromo-2,3-difluorophenol (**25**) were chosen as starting materials to access the target molecules **22** and **23**, respectively. The Negishi cross-coupling of 4-bromo-2-fluorophenol was not investigated as 3-fluorotyrosine is commercially available.

There is precedence for the use of 2-, 3- or 4-halo-phenols as substrates in the Pd-catalysed Negishi cross-coupling reaction.[48] The reaction was found to be tolerant to the acidity of a phenol hydroxyl group. However, comparatively diminished yields were observed.[48] Fluorine substitution is known to increase acidity of the phenol hydroxyl group and the tolerance of the Negishi reaction to fluorinated iodo-phenols is

unknown.[100] Therefore, a more tentative strategy that utilises a hydroxyl protecting group was chosen. As the β-iodo amino acid being used is **1** an appropriate protection of the hydroxyl group would match either the *N*-Boc or OBzl so that it could be removed in the same step and truncate the route to the unprotected amino acids needed to supplement into the growth medium of *Bacillus* sp. CS93. OBzl protection of the phenol hydroxyl was chosen and protected products (**26** and **27**) were obtained in good yield and no efforts were made to further optimise this reaction (**Figure 2.20**).



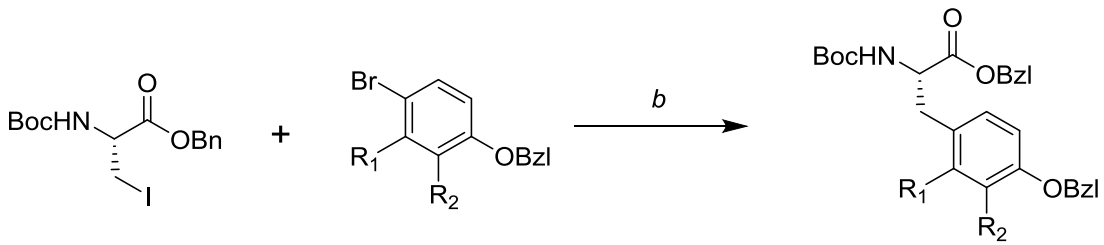
24: R_1=F, R_2=H
25: R_1, R_2= F

26: R_1=F, R_2=H (**64%***)
27: R_1, R_2= F (**77%***)

Figure 2.20: Synthesis of **26** and **27**. *i*: K_2CO_3 (1.20 eq), benzyl bromide (1.10 eq), DMF, RT, 18 h. *Denotes isolated yield of pure product.

2.5.3 Synthesis of Fluoro-Tyrosines

A Negishi reaction was carried out using **26** and **27** with our best previous conditions (*b* from **Section 2.3.3**, **Scheme 2.19**).



1

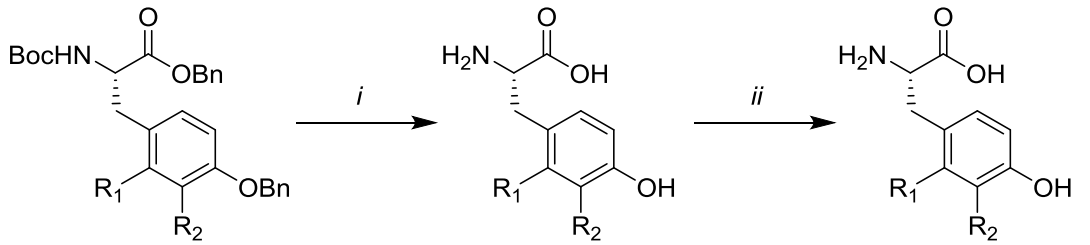
26: R_1=F, R_2=H
27: R_1, R_2= F

28: R_1=F, R_2=H (**72%***)
29: R_1, R_2= F (**60%***)

Scheme 2.19: Synthesis of **28** and **29**. *b*: Zn (4.00 eq), 100 °C, 20 min, *in vacuo*; I_2 (cat.), DMF, 70 °C, 20 min, argon; **1**, 50 °C, 20 min; **26** or **27** (1.00 eq), $Pd_2(dba)_3$ (0.02 eq), SPhos (0.07 eq), DMF, 50 °C, 5 h; RT, 18 h. *Denotes isolated yield of pure product.

The Pd-catalysed Negishi cross-coupling reaction gave protected product (**28** and **29**) in good yield. The fluoro-tyrosines were subsequently deprotected by Dr. O'Connor at

University College Dublin to give target molecules **22** and **23** in 71% and 57% yields, respectively (**Scheme 2.20**).



28: R_1=F, R_2=H
29: R_1, R_2= F

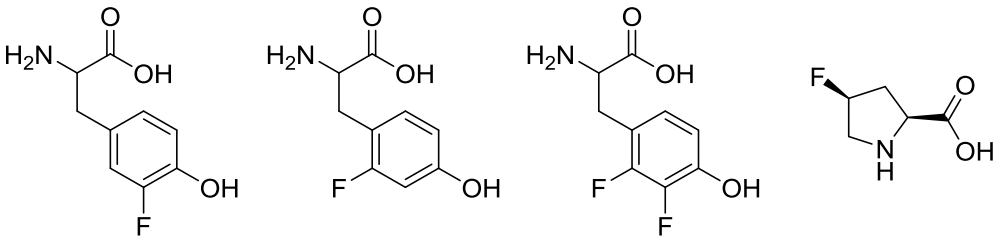
22: R_1=F, R_2=H (**71%***)
23: R_1, R_2= F (**57%***)

Scheme 2.20: *i*: HCl(aq), MeOH ,RT, 4 h.; *ii*: H_2, Pd(cat), MeOH.** *Denotes isolated product after HPLC purification. **If steps *i* and *ii* reveal methylation of the carboxylic acid a subsequent basic step is undertaken: KOH, MeOH, H_2O, RT, 24 h.

2.5.4 Biosynthesis of Fluoro-Tyrosine Lipopeptides: Iturin A and Fengycin

Below is a summary of the biosynthetic work that was carried out by Dr. O'Connor at University College Dublin and Durham University. This work has also been published in Amino Acids.[20]

Amino acids **22** and **23** as well as 3-fluorotyrosine and *cis*-fluoroproline (**30**)* (**Figure 2.21**) were supplemented into the growth medium of *Bascillus* sp. CS93 and extracts from the cultures were analysed for fluorinated lipopeptide content.



3-fluorotyrosine | **22** 2-fluorotyrosine | **23** 2,3-difluorotyrosine | **30** cis-fluoroproline

Figure 2.21: The amino acids supplemented to the growth medium of *Bascillus* sp. CS93.

* The synthesis of **30** is described in **Appendix 1**.

The analysis of culture extracts showed that fluorinated tyrosine amino acids were incorporated into iturin A and another lipopeptide: Fengycin (**Figure 2.22**) to differing extents (**Table 2.6**).

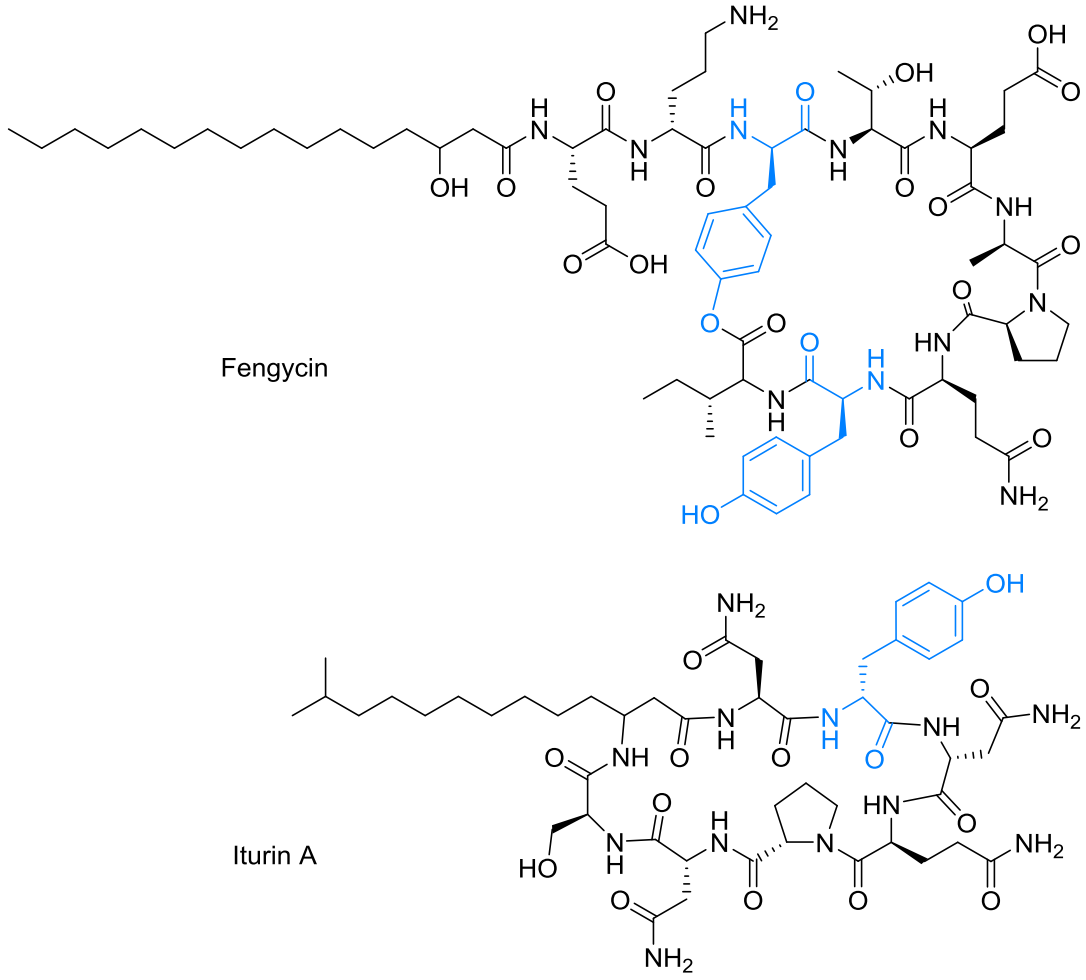


Figure 2.22: Structures of C-16 fengycin and C-14 iturin A. Tyrosines are labelled in blue.

Entry	Compound No.	Amino Acid	Incorporation	
			Iturin A	Fengycin
1	-	3-fluorotyrosine	✓	✓
2	**22**	2-fluorotyrosine	✗	✗
3	**23**	2,3-difluorotyrosine	✗	✓
4	**30**	cis-fluoroproline	✗	✗

Table 2.6: Observed incorporation of fluorinated tyrosines into lipopeptides.

Purified fluoro-iturin and mono-fluoro-fengycin were tested for anti-fungal activity in a bioassay. The results showed similar but not improved potency compared to the natural products: Iturin A and a mixture of iturin/fluoroiturin showed a minimum inhibitory concentration (MIC) of 32 μg/mL against the fungus: Trichophyton rubrum. Whereas, Fengycin (C-16 and C-17) and the 3-flourotyrosine or 2,3-difluorotyrosine containing fluorinated derivatives exhibited no inhibition (at the highest concentration tested: 64 μg/mL).[20]

2.5.5 Fluorinated Tyrosine Conclusions

Two fluorinated tyrosine amino acids (**22** and **23**) have been efficiently synthesised *via* Pd-catalysed Negishi cross-coupling. The synthetic route is a major improvement on the synthetic routes previously described.[19,98,99] The fluorinated amino acids were amenable (to varying extents) to incorporation *via* biosynthetic methods into natural products: fengycin and iturin A.

2.6 General Conclusions on Heteroaromatic Synthesis *via* Pd-Catalysed Cross-Coupling

The Negishi cross-coupling strategy is an extremely useful "tool" in the synthesis of a variety of unnatural aromatic amino acids from simple building blocks (**1** or **9**). A series of previously difficult to access (*via* Negishi cross-coupling) furyl- and thienyl- amino acids (X=O and X=S, respectively) as well as previously undescribed (*via* Negishi cross-coupling) fluoro-tyrosines were synthesised. These amino acids were synthesised with protecting group (P_1 and P_2) strategies to facilitate their translation into bioactive products. The *N*-Boc (P_2), OBzl (P_1) strategy proved successful in the synthesis of fluorotyrosines and 2,3-difluorotyrosine (**23**) was incorporated into the fengycin via biosynthesis (**Section 2.4.5**). In the synthesis of furylalanines, the *N*-Boc (P_2), OBzl (P_1) strategy proved high yielding, yet the 2-furyalanine derivative (**19**) was not stable to deprotection *via* hydrogenation. However, the lower yielding *N*-Fmoc (P_2) O*t*Bu (P_1) has enabled the 2-furyalanine (**13**) derivative to be incorporated into a DKP scaffold that will be tested for inhibitory activity against CCL2 induced chemotaxis (**Chapter 4**).

2.7 References

1. A. K. Sato, M. Viswanathan, R. B. Kent and C. R. Wood, *Curr. Opin. Biotechnol.*, 2006, **17**, 638-642.
2. S. A. Marshall, G. A. Lazar, A. J. Chirino and J. R. Desjarlais, *Drug Discov. Today*, 2003, **8**, 212-221.
3. D. Schwarzer, C. Ludwig, I. V. Thiel and H. D. Mootz, *Biochemistry*, 2012, **51**, 233-242.
4. K.-H. Lee, C. Cabello, L. Hemmingsen, E. N. G. Marsh and V. L. Pecoraro, *Angew. Chem.-Int. Edit.*, 2006, **45**, 2864-2868.
5. J. S. Thorson, E. Chapman, E. C. Murphy, P. G. Schultz and J. K. Judice, *J. Am. Chem. Soc.*, 1995, **117**, 1157-1158.
6. S. J. Miller, H. E. Blackwell and R. H. Grubbs, *J. Am. Chem. Soc.*, 1996, **118**, 9606-9614.
7. M. J. P. de Vega, M. I. Garcia-Aranda and R. Gonzalez-Muniz, *Med. Res. Rev.*, 2011, **31**, 677-715.
8. A. S. Hudson, A. Hoose, C. R. Coxon, G. Sandford and S. L. Cobb, *Tetrahedron Lett.*, 2013, **54**, 4865-4867.
9. C. Frieden, S. D. Hoeltzli and J. G. Bann, *Methods Enzymol.*, 2004, **380**, 400-415.
10. M. Suchanek, A. Radzikowska and C. Thiele, *Nat. Methods*, 2005, **2**, 261-267.
11. J. L. Kitevski-LeBlanc and R. S. Prosser, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2012, **62**, 1-33.
12. N. K. Mishra, A. K. Urick, S. W. J. Ember, E. Schönbrunn and W. C. Pomerantz, *ACS Chem. Biol.*, 2014, **9**, 2755-2760.
13. C. L. Weeks, A. Polishchuk, Z. Getahun, W. F. Degrado and T. G. Spiro, *J. Raman Spectrosc.*, 2008, **39**, 1606-1613.
14. P. Marek, R. Gupta and D. P. Raleigh, *Chembiochem*, 2008, **9**, 1372-1374.
15. A. Perdih and M. S. Dolenc, *Curr. Org. Chem.*, 2011, **15**, 3750-3799.
16. A. E. Pojitkov, E. N. Efremenko and S. D. Varfolomeev, *J. Mol. Catal. B: Enzym.*, 2000, **10**, 47-55.
17. D. B. Cowie and G. N. Cohen, *Biochim. Biophys. Acta*, 1957, **26**, 252-261.
18. N. Blatter, A. Prokup, A. Deiters and A. Marx, *Chembiochem*, 2014, **15**, 1735-1737.
19. B. Brooks, R. S. Phillips and W. F. Benisek, *Biochemistry*, 1998, **37**, 9738-9742.
20. N. K. O'Connor, A. S. Hudson, S. L. Cobb, D. O'Neil, J. Robertson, V. Duncan and C. D. Murphy, *Amino Acids*, 2014, **46**, 2745-2752.
21. A. M. Spokoyny, Y. Zou, J. J. Ling, H. Yu, Y.-S. Lin and B. L. Pentelute, *J. Am. Chem. Soc.*, 2013, **135**, 5946-5949.
22. M. Saleki, N. Colgin, J. A. Kirby, S. L. Cobb and S. Ali, *MedChemComm*, 2013, **4**, 860-864.
23. M. G. Hoesl and N. Budisa, *Curr. Opin. Biotechnol.*, 2012, **23**, 751-757.
24. C. C. Liu and P. G. Schultz, *Annu. Rev. Biochem.*, 2010, **79**, 413-444.
25. N. Voloshchuk and J. K. Montclare, *Mol. Biosyst.*, 2010, **6**, 65-80.
26. K. Harada, *Nature*, 1963, **200**, 1201.
27. M. S. Iyer, K. M. Gigstad, N. D. Namdev and M. Lipton, *Amino Acids*, 1996, **11**, 259-268.
28. X. H. Cai and B. Xie, *Arkivoc*, 2014, 205-248.
29. Y.-L. Liu and J. Zhou, *Chem. Comm.*, 2013, **49**, 4421-4423.

30. T. Ooi, Y. Uematsu, J. Fujimoto, K. Fukumoto and K. Maruoka, *Tetrahedron Lett.*, 2007, **48**, 1337-1340.
31. P. N. Collier, A. D. Campbell, I. Patel, T. M. Raynham and R. J. K. Taylor, *J. Org. Chem.*, 2002, **67**, 1802-1815.
32. A. J. Ross, H. L. Lang and R. F. W. Jackson, *J. Org. Chem.*, 2010, **75**, 245-248.
33. N. Colgin, T. Flinn and S. L. Cobb, *Org. Biomol. Chem.*, 2011, **9**, 1864-1870.
34. M. Suhartono, M. Weidlich, T. Stein, M. Karas, G. Dürner and M. W. Göbel, *Eur. J. Org. Chem.*, 2008, **2008**, 1608-1614.
35. W.-C. Shieh, S. Xue, N. Reel, R. Wu, J. Fitt and O. Repič, *Tetrahedron: Asymmetry*, 2001, **12**, 2421-2425.
36. H. J. Kreuzfeld, C. Döbler, U. Schmidt and H. W. Krause, *Amino Acids*, 1996, **11**, 269-282.
37. B. D. Vineyard, W. S. Knowles, M. J. Sabacky, G. L. Bachman and D. J. Weinkauff, *J. Am. Chem. Soc.*, 1977, **99**, 5946-5952.
38. Y. Belokon, V. Maleev, T. Savel'eva, M. Moskalenko, D. Pripadchev, V. Khrustalev and A. Saghiyan, *Amino Acids*, 2010, **39**, 1171-1176.
39. V. A. Soloshonok, C. Cai, T. Yamada, H. Ueki, Y. Ohfune and V. J. Hruby, *J. Am. Chem. Soc.*, 2005, **127**, 15296-15303.
40. T. K. Ellis, C. H. Martin, G. M. Tsai, H. Ueki and V. A. Soloshonok, *J. Org. Chem.*, 2003, **68**, 6208-6214.
41. J. Zhou and G. C. Fu, *J. Am. Chem. Soc.*, 2003, **125**, 12527-12530.
42. E. Negishi, *Acc. Chem. Res.*, 1982, **15**, 340-348.
43. G. Lessene, *Aust. J. Chem.*, 2004, **57**, 107-107.
44. R. F. W. Jackson, N. Wishart, A. Wood, K. James and M. J. Wythes, *J. Org. Chem.*, 1992, **57**, 3397-3404.
45. J. M. Brown and N. A. Cooley, *Chem. Rev.*, 1988, **88**, 1031-1046.
46. V. B. Phapale and D. J. Cardenas, *Chem. Soc. Rev.*, 2009, **38**, 1598-1607.
47. E. I. Negishi, *Angew. Chem.-Int. Edit.*, 2011, **50**, 6738-6764.
48. R. F. W. Jackson, I. Rilatt and P. J. Murray, *Org. Biomol. Chem.*, 2004, **2**, 110-113.
49. G. Manolikakes, C. Muñoz Hernandez, M. A. Schade, A. Metzger and P. Knochel, *J. Org. Chem.*, 2008, **73**, 8422-8436.
50. I. Rilatt, L. Caggiano and R. F. W. Jackson, *Synlett*, 2005, 2701-2719.
51. R. F. W. Jackson, M. J. Wythes and A. Wood, *Tetrahedron Lett.*, 1989, **30**, 5941-5944.
52. C. L. Oswald, T. Carrillo-Márquez, L. Caggiano and R. F. W. Jackson, *Tetrahedron*, 2008, **64**, 681-687.
53. C. Knox, V. Law, T. Jewison, P. Liu, S. Ly, A. Frolkis, A. Pon, K. Branco, C. Mak, V. Neveu, Y. Djoumbou, R. Eisner, A. C. Guo, D. Cheng, S. Shrivastava, D. Tzur, B. Gautam, M. Hassanali, P. Stothard, Z. Chang and J. Woolsey, *Furosemide*, DrugBank, 2005.
54. A. Pan, I. G. Stiell, R. Dionne and J. Maloney, *Emerg. Med. J.*, 2015, **32**, 36-77.
55. C. Knox, V. Law, T. Jewison, P. Liu, S. Ly, A. Frolkis, A. Pon, K. Branco, C. Mak, V. Neveu, Y. Djoumbou, R. Eisner, A. C. Guo, D. Cheng, S. Shrivastava, D. Tzur, B. Gautam, M. Hassanali, P. Stothard, Z. Chang and J. Woolsey, *Ethylmethylthiambutene*, DrugBank, 2007.
56. *International Pat.*, WO2003097608-A2, 2003.
57. S. A. Salmon, J. L. Watts and R. J. Yancey, *J. Vet. Diagn. Invest.*, 1996, **8**, 332-336.

58. S. A. Brown, B. J. Hanson, A. Minot, L. Millerioux, P. J. Hamlow, V. L. Hubbard, J. K. Callahan and F. M. Kausche, *J. Vet. Pharmacol. Ther.*, 1999, **22**, 35-40.
59. D. Wile, *Ann. Clin. Biochem.*, 2012, **49**, 419-431.
60. A. F. Green, *Br. J. Pharmacol. Chemother.*, 1953, **8**, 2-9.
61. M. Hilker, C. Häberlein, U. Trauer, M. Bünnige, M.-O. Vicentini and S. Schulz, *Chembiochem*, 2010, **11**, 1720-1726.
62. L. P. Partida-Martinez, C. Flores de Looß, K. Ishida, M. Ishida, M. Roth, K. Buder and C. Hertweck, *Appl. Envirom. Microb.*, 2007, **73**, 793-797.
63. U. Schmidt, A. Lieberknecht and J. Wild, *Synthesis*, 1984, 53-60.
64. A. S. K. Hashmi, P. Haufe, C. Schmid, A. Rivas Nass and W. Frey, *Chem. -Eur. J.*, 2006, **12**, 5376-5382.
65. Y. Koseki, H. Yamada and T. Usuki, *Tetrahedron: Asymmetry*, 2011, **22**, 580-586.
66. R. F. W. Jackson, R. J. Moore, C. S. Dexter, J. Elliott and C. E. Mowbray, *J. Org. Chem.*, 1998, **63**, 7875-7884.
67. J. B. Tuttle, J. M. Azzarelli, B. M. Bechle, A. B. Dounay, E. Evrard, X. Gan, S. Ghosh, J. Henderson, J.-Y. Kim, V. D. Parikh and P. R. Verhoest, *Tetrahedron Lett.*, 2011, **52**, 5211-5213.
68. M. Tanaka, H. Hikawa and Y. Yokoyama, *Tetrahedron*, 2011, **67**, 5897-5901.
69. M. F. Eerland and C. Hedberg, *J. Org. Chem.*, 2012, **77**, 2047-2052.
70. H. J. C. Deboves, C. A. G. N. Montalbetti and R. F. W. Jackson, *J. Chem. Soc., Perkin Trans. 1*, 2001, 1876-1884.
71. D. Fillion, M. Deraët, B. J. Holleran and E. Escher, *J. Med. Chem.*, 2006, **49**, 2200-2209.
72. S. Tabanella, I. Valancogne and R. F. W. Jackson, *Org. Biomol. Chem.*, 2003, **1**, 4254-4261.
73. M. S. Smyth and T. R. Burke Jr, *Tetrahedron Lett.*, 1994, **35**, 551-554.
74. M. Ousmer, V. Boucard, N. Lubin-Germain, J. Uziel and J. Augé, *Eur. J. Org. Chem.*, 2006, **2006**, 1216-1221.
75. C. Hoppmann, U. Alexiev, E. Irran and K. Rück-Braun, *Tetrahedron Lett.*, 2013, **54**, 4585-4587.
76. D. A. Culkin and J. F. Hartwig, *Organometallics*, 2004, **23**, 3398-3416.
77. T. E. Barder, S. D. Walker, J. R. Martinelli and S. L. Buchwald, *J. Am. Chem. Soc.*, 2005, **127**, 4685-4696.
78. R. Martin and S. L. Buchwald, *Acc. Chem. Res.*, 2008, **41**, 1461-1473.
79. F. Paul, J. Patt and J. F. Hartwig, *J. Am. Chem. Soc.*, 1994, **116**, 5969-5970.
80. A. A. C. Braga, N. H. Morgon, G. Ujaque and F. Maseras, *J. Am. Chem. Soc.*, 2005, **127**, 9298-9307.
81. V. F. Slagt, A. H. M. de Vries, J. G. de Vries and R. M. Kellogg, *Org. Process Res. Dev.*, 2009, **14**, 30-47.
82. A. M. Webster, C. R. Coxon, A. M. Kenwright, G. Sandford and S. L. Cobb, *Tetrahedron*, 2014, **70**, 4661-4667.
83. A. Schulz, A. Busmann, E. Kluver, M. Schnebel and K. Adermann, *Protein Pept. Lett.*, 2004, **11**, 601-606.
84. A. S. Hudson, L. Caron, N. Colgin and S. L. Cobb, *Amino Acids*, 2015, **47**, 779-785.
85. J. T. Gerig, *Prog. Nucl. Magn. Reson. Spectrosc.*, 1994, **26**, 293-370.
86. R. Filler and R. Saha, *Future Med. Chem.*, 2009, **1**, 777-791.

87. J. Wang, M. Sánchez-Roselló, J. L. Aceña, C. del Pozo, A. E. Sorochinsky, S. Fustero, V. A. Soloshonok and H. Liu, *Chem. Rev.*, 2013, **114**, 2432-2506.
88. C. Murphy, B. Clark and J. Amadio, *Appl. Microbiol. Biotechnol.*, 2009, **84**, 617-629.
89. J. Pepin, C. Guern, F. Milord and P. J. Schechter, *Lancet*, 1987, **330**, 1431-1433.
90. B. Moser and P. Loetscher, *Nat. Immunol.*, 2001, **2**, 123-128.
91. C. Gerard and B. J. Rollins, *Nat. Immunol.*, 2001, **2**, 108-115.
92. H. J. Dyson and P. E. Wright, *Elucidation of the protein folding landscape by NMR*, Elsevier Academic Press Inc, San Diego, 2005.
93. N. K. O'Connor, D. K. Rai, B. R. Clark and C. D. Murphy, *J. Fluor. Chem.*, 2012, **143**, 210-215.
94. S. K. Holmgren, K. M. Taylor, L. E. Bretscher and R. T. Raines, *Nature*, 1998, **392**, 666-667.
95. A. L. Serrano, T. Troxler, M. J. Tucker and F. Gai, *Chem. Phys. Lett.*, 2010, **487**, 303-306.
96. S. Moran, D. K. Rai, B. R. Clark and C. D. Murphy, *Org. Biomol. Chem.*, 2009, **7**, 644-646.
97. L. Volpon, F. Besson and J.-M. Lancelin, *Eur. J. Biochem.*, 1999, **264**, 200-210.
98. R. L. VonTersch, F. Secundo, R. S. Phillips and M. G. Newton, *Biomedical Frontiers of Fluorine Chemistry*, American Chemical Society, 1996.
99. M. Monclus, C. Masson and A. Luxen, *J. Fluor. Chem.*, 1995, **70**, 39-43.
100. M. D. Liptak, K. C. Gross, P. G. Seybold, S. Feldgus and G. C. Shields, *J. Am. Chem. Soc.*, 2002, **124**, 6421-6427.

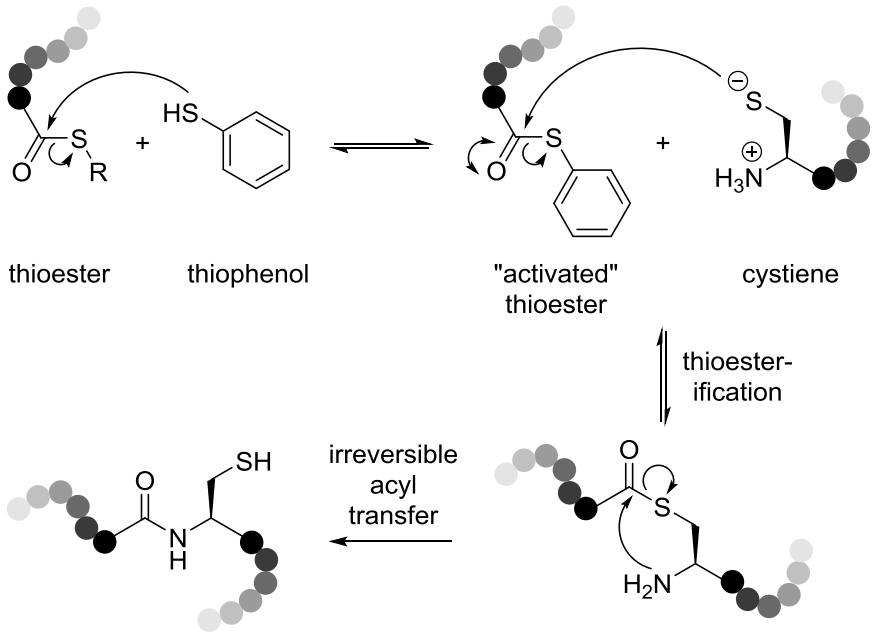
Chapter 3 : The Synthesis and Applications of Thiol Containing Amino Acids in NCL

3.1 Thiol Containing Amino Acids

3.1.1 Native Chemical Ligation (NCL)

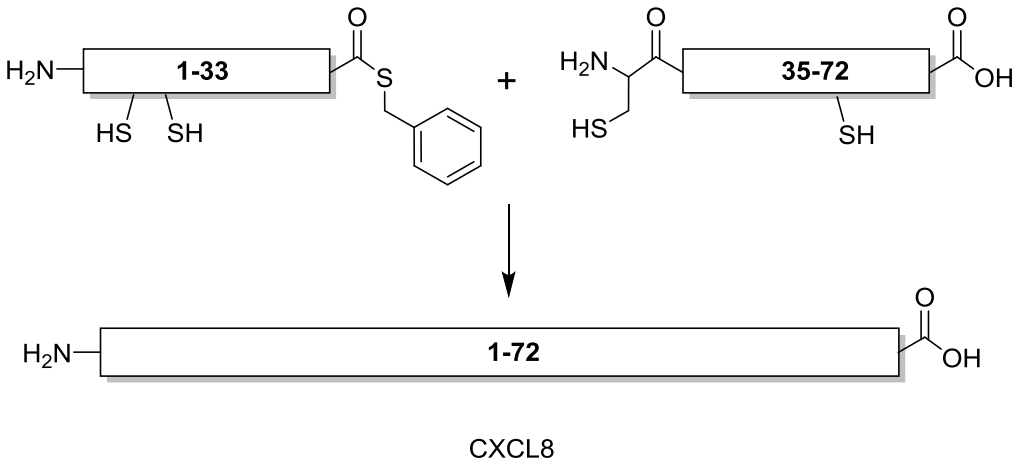
Since its discovery native chemical ligation (NCL) has enabled routes to the total chemical synthesis of numerous proteins[1] and the semi-synthesis of even larger biological constructs.[2,3] NCL originates from a reaction that has been known since 1956 where Wieland *et al.* synthesised a valine-cysteine dipeptide in aqueous buffer from a valine-thioester and a native cysteine.[4] In 1994 Kent *et al.* were able to extend this reaction to form a peptide chain from two fully deprotected peptide segments.[1]

NCL relies on the chemoselective coupling of two unprotected peptides in a neutral aqueous solution, one bearing a C-terminal thioester and the other an N-terminal cysteine (**Scheme 3.1**). NCL is carried out under denaturing conditions (to inhibit the formation of secondary structures that lower peptide solubility) and typically in the presence of a thiol catalyst. The thiol catalyst (e.g. thiophenol) is added to the reaction so that a more activated thioester is formed *in situ.* [5]



Scheme 3.1: The general mechanism for NCL: The first step involves formation of the "activated" thioester on the C-terminal segment. This is subsequently attacked by the N-terminal cysteine segment. The final step is the irreversible acyl-transfer yielding a native peptide bond and a cysteine thiol.

One of the key advantages of NCL is that the reaction can be carried out on unprotected peptides and peptides that contain more than one cysteine residue. An example of this can be seen in the preparation of Human interleukin-8 (IL-8 or CXCL8). CXCL8 is a 72-amino acid chemokine that was synthesised from two peptide segments (of 33 and 39 residues in length) using NCL (**Scheme 3.2**).[1] The reaction proceeded regio-selectively to give the desired product, where the N-terminal cysteine had been utilised in preference over the internal cysteine in the IL-8 (35-72) fragment. Tolerance to 18 of the 20 total natural amino acids was shown as well as the lack of interference from internal other cysteines.[1] The resulting protein was purified and oxidised to give a folded protein that displayed the same physical and biological properties as a reference sample of CXCL8.[1]

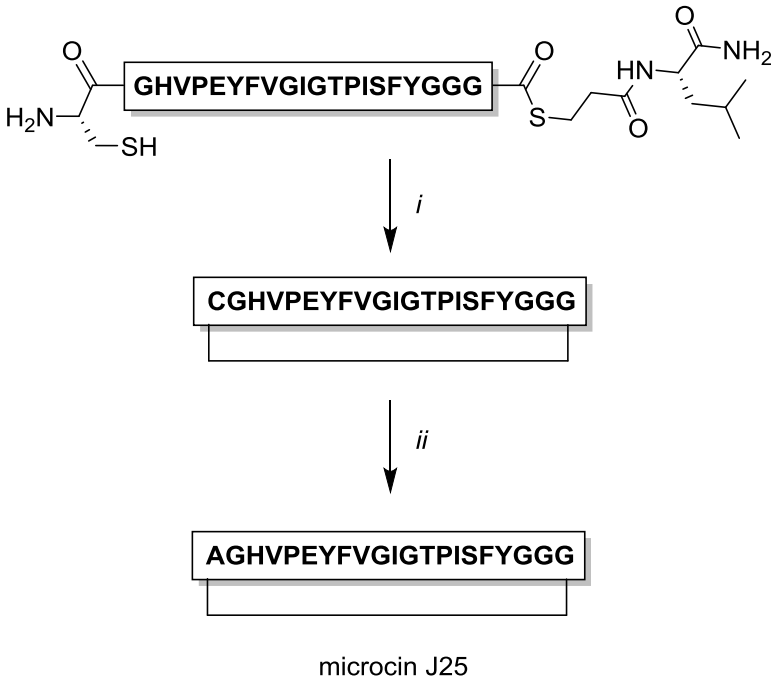


Scheme 3.2: The deprotected and purified peptide segments are reacted in a NCL reaction. Segment 1-33 contains a thioester moiety and 35-72 contains an N-terminal cysteine.

The chemo-selectivity observed in the NCL reaction (i.e. cysteine over lysine or serine nucleophiles) is due to the chemical properties of thioesters. Thioesters are regarded as being more stable to hydroxide-catalysed hydrolysis than oxoesters and more reactive to thiolysis and aminolysis.[5] Thus, unprotected peptides can be ligated *via* the N-terminal cysteine without the formation of by-products formed from other nucleophilic side chains in the peptide fragments. Generally, NCL has a wide scope and gives good yields with low levels of epimerisation. NCL has become a keystone in the chemical synthesis of proteins. Thus, numerous approaches to overcome its limitations are being investigated.[5,6]

3.1.3 Native Chemical Ligation Limitations

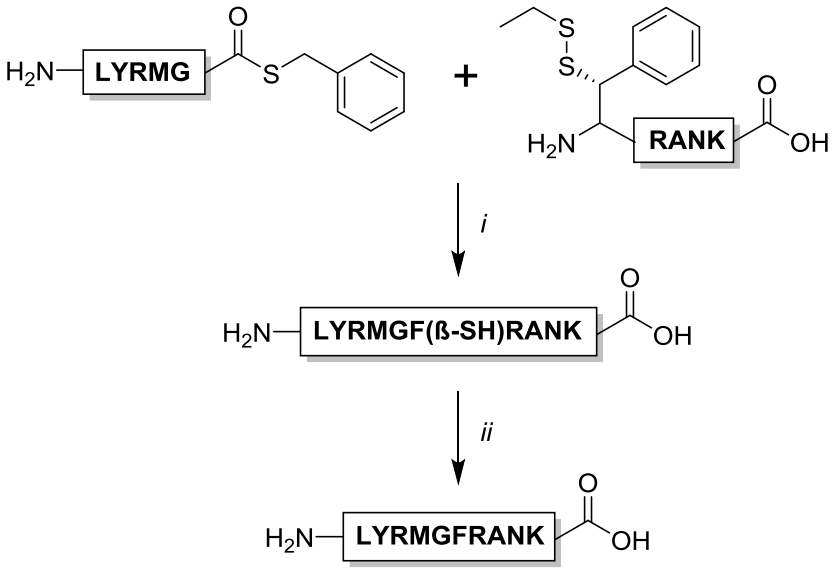
A major limitation of NCL is that the protein must not only contain a cysteine residue but also one in a preferred central site for ligation.[5] However, the first step to expand NCL from strictly Xaa-Cys junctions was to utilise a post-ligation desulfurisation reaction of cysteine to form alanine. This reaction was introduced by Yan and Dawson in the total chemical syntheses of various cysteine-free peptides. For example, the cyclic antibiotic microcin J25.[7] An alanine was replaced by a cysteine and used in ligation. Upon intramolecular ligation the cysteine is specifically desulfurised by H_2/metal catalysts (primarily Pd/Al_2O_3) to yield the native alanine (Xaa-Ala) (**Scheme 3.3**).[7]



Scheme 3.3: The NCL reaction (*i*) and the subsequent desulfurisation reaction (*ii*) that yield a peptide: microcin J25 with a native amide bond between A and G residues. *i*: 0.1M tris-HCl, 6 M Gdn, benzyl mercaptan, thiophenol pH 8.5. *ii*: H_2, Pd/Al_2O_3, AcOH.[7]

Advances in the field have enabled tolerance of other cysteines *via* orthogonal protection (e.g. with Acm)[8] and metal-free strategies have been developed using a radical initiator as well as a trialkylphosphine.[9] Perhaps the most important advancement in this field comes from the use of unnatural amino acids that can easily be inserted into peptide chains with SPPS.

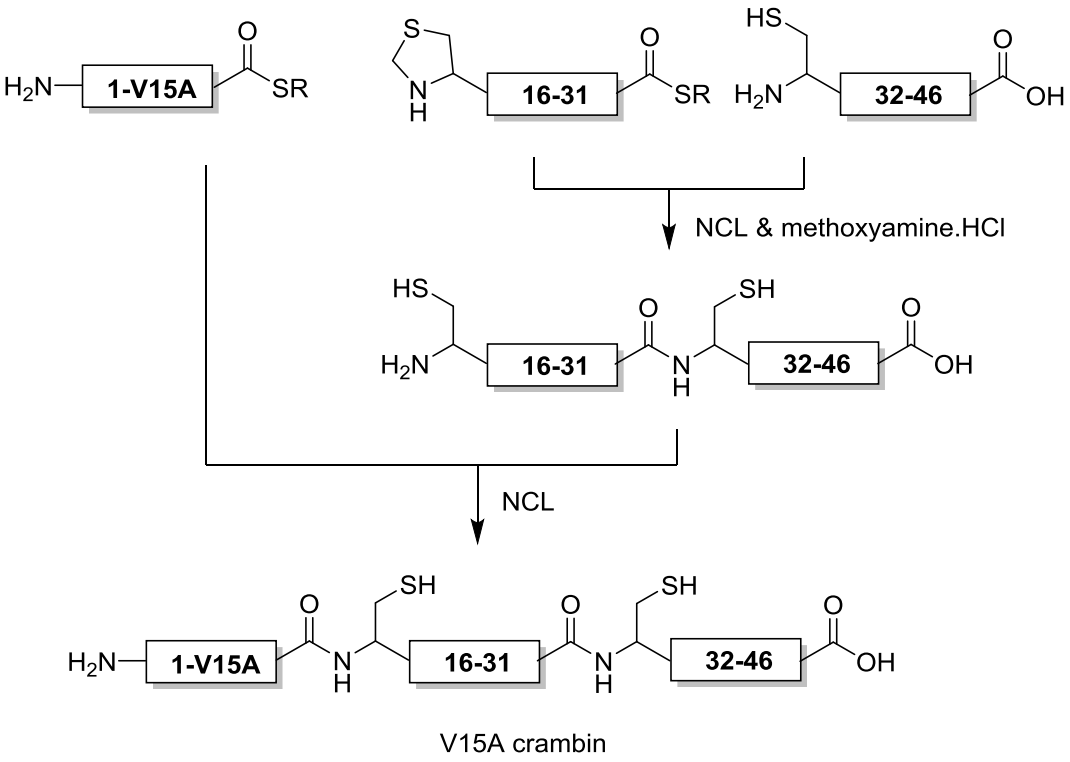
The first unnatural thiol amino acid to be used in NCL was penicillamine.[6] N-terminal penicillamine was used in a NCL and desulfurisation of the product yielded a Xaa-Val amide in the peptide sequence. Since these early investigations a variety of β-thiol containing amino acids have been incorporated into peptides and effectively used in the joint NCL/desulfurisation combination reactions to yield Xaa-Leu,[10] Xaa-Pro[11] and Xaa-Phe[12](**Scheme 3.4**) junctions.



Scheme 3.4: The NCL reaction (*i*) and the subsequent desulfurisation reaction (*ii*) that yield a peptide (LYRMGFRANK) with a native amide bond between G and F residues. 0.1M tris-HCl, MeSNa, TCEP, pH 8. *ii*: 0.1M phosphate buffer, $NaCl_2.6H_2O$, $NaBH_4$, pH 7.

3.1.3 Multiple Ligation Strategies: N-terminal Cysteine Protection

When a stepwise approach to ligations of numerous segments is undertaken, orthogonal yet conventional protecting group strategies must be used to keep terminal cysteines masked until they are needed. An example is the total one-pot synthesis of (V15A)crambin (**Scheme 3.5**).[13] The first step in this strategy is the ligation of the Thz protected (terminal cysteine) peptide segment **16-31** to segment **32-46**. The Thz protection is then removed (with methoxyamine.HCl) and a further ligation of this product to the last segment **1-V15A**, yields the full peptide. The reactions are carried out in one-pot without intermediate purification steps.[13]

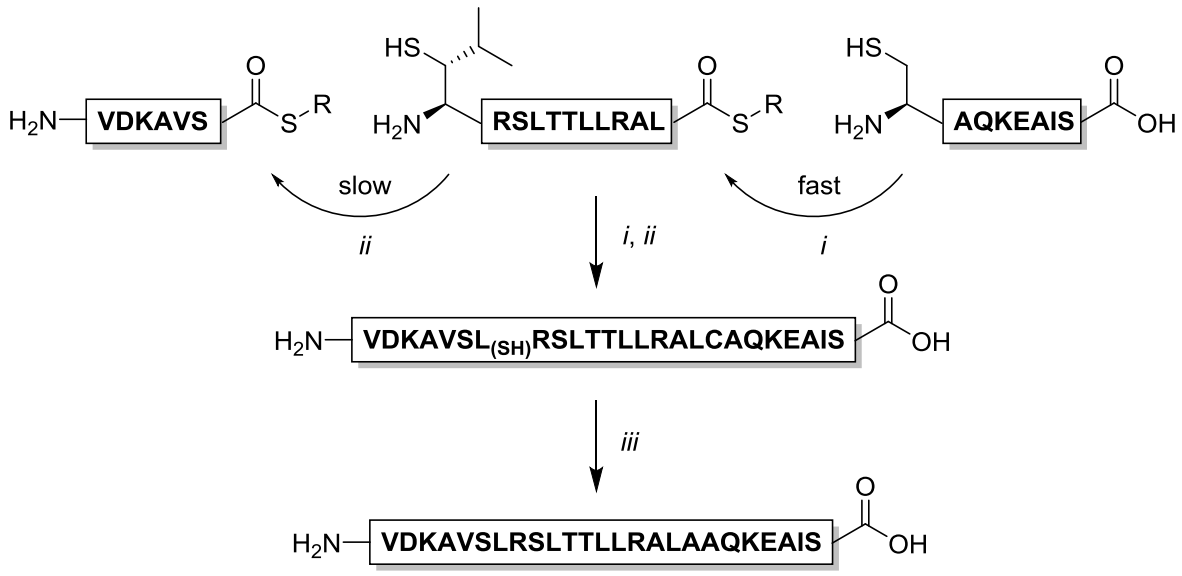


Scheme 3.5: The one-pot synthesis of (V15A)crambin from three peptide segments (1-V15A, 16-31 and 32-46).

3.1.4 Multiple Ligation Strategies: Kinetic Native Chemical Ligation

Another major advancement in the approach to multiple NCL strategies is to tailor the reaction kinetics so that multiple peptide segments can be ligated together in a controllable and sequential manner. For example, when a normal "unreactive" alkyl thioester and a more reactive (pre-formed) aryl thioester are competitively ligated, the aryl thioester successfully outcompetes the alkyl thiol (until exogenous thiol catalyst is added).[14]

New avenues for the development of kinetically controlled NCL reactions have recently opened up due to the application of unnatural N-terminal β-thiol amino acids. The Danishefsky group was able to utilise a noted difference in reactivity (towards NCL) in two N-terminal β-thiol amino acids (cystiene and (*R*,*R*)mercaptoleucine) to perform a kinetically controlled ligation followed by desulfurisation to yield an alanine and leucine residue at the ligation junctions, respectively (**Scheme 3.6**).



Scheme 3.6: A two-step kinetically controlled NCL reaction. *i*: 6 M Gn•HCl, 100 mM NaH_2PO_4, 50 mM TCEP, pH 7.5, 0.5 h. *ii*: 6 M Gdn HCl, 100 mM NaH_2PO_4, 50 mM TCEP, pH 7.5, 0.5 h; MESNa, H_2O:MeCN (1:1), 1 min. *iii*: TCEP, VA-044, *t*BuSH, 1h.

Moreover, significant differences in NCL reactivity were also noted between (*R*,*S*) and (*R*,*R*)-β-mercaptoleucine. The ligation yields for the (*R*,*S*) and (*R*,*R*) diastereoisomers (under the same conditions) were 75% or 12%, respectively.[15] Following this, the Danishefsky group showed the stereochemical effect to be more pronounced on *cis*- and *trans*- thio-proline amino acids. Under the same conditions the *cis*- and *trans*-diastereoisomers showed <5% and 85% conversion to ligated product, respectively.[11]

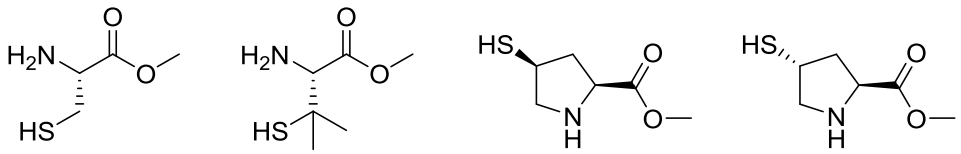
With β-thiol amino acids exhibiting major differences in reactivity (yield and/or reaction time) in specific NCL conditions and changes to just one stereo-centre leading to a large effect[11] it is surprising that extensive biophysical studies are lacking in this area. Thus, more information on the nature of the thio- amino acids would be beneficial as detailed insights into their NCL reactions are needed in order to expand the area of kinetically controlled NCL reactions. A starting point to achieve this understanding is to measure the pK_as of a series of thiol amino acids.

3.2 Syntheses of Thiol Containing Amino Acids

3.2.1 Synthesis of a Small Library of Thiol Containing Amino Acids

Many thio- amino acids are commercially available (e.g. Cys and Pen, giving Ala[7] or Val[6] ligations, respectively) or routes to their synthesis are known. Therefore, ligations at various sites (e.g. Phe,[12] Thr,[16] Leu,[10] and Lys[17]) has been accomplished. However, many of the amino acid synthesis approaches reported are challenging and give poor control of stereochemistry. Given this, the development of more efficient syntheses of β-thiol amino acids that also give good stereo- and regio-control would be of benefit to this area of research.

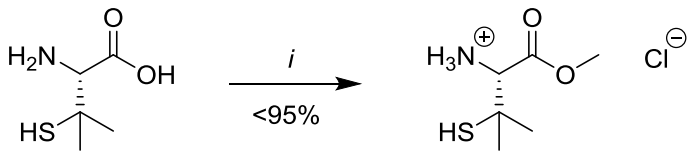
In our study four thiol containing amino acids (compounds **31** – **34** in **Figure 3.1**) were chosen for the pKa studies. All amino acids contain a methyl ester substituent on the carbocylic acid, a free amino group and importantly a free thiol.



31 **32** **33** **34**

Figure 3.1: Target library of four thiol containing amino acids, **31** - **34**.

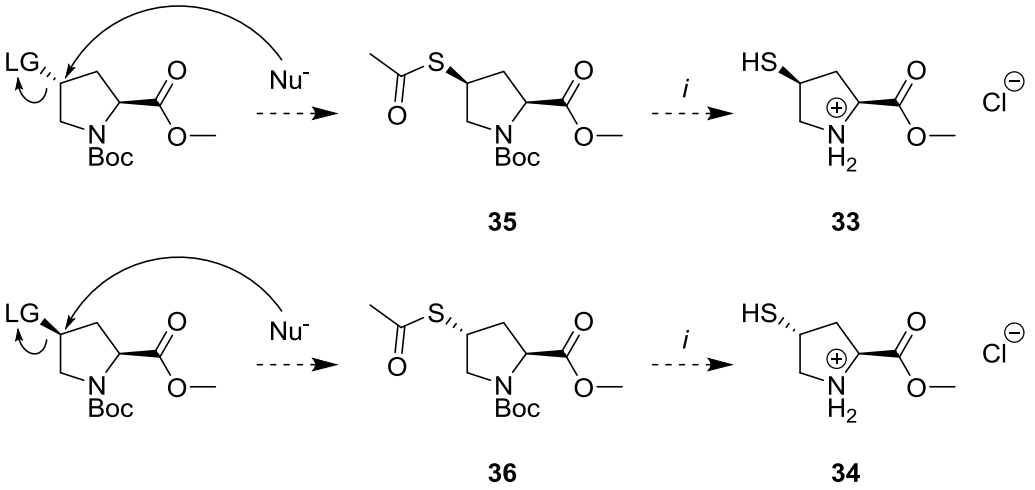
Compound **31** is commercially available (Sigma) as a hydrochloride salt and it is inexpensive. Compound **32** was readily synthesised as a hydrochloride salt from L-penicillamine (H-Pen-OH, Bachem) as in **Scheme 3.7** below.



L-penicillamine **32**

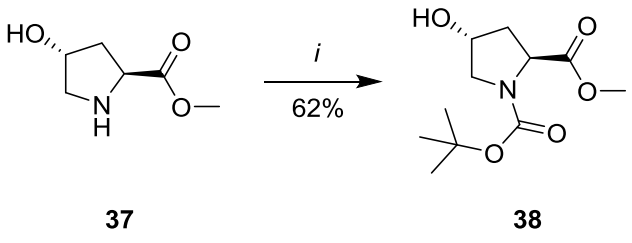
Scheme 3.7: Synthesis of **32**. *i*: L-penicillamine, thionyl chloride, methanol, reflux 48 h.

The majority of published syntheses of unnatural thiol-amino acids is focused around nucleophilic substitution chemistry on the corresponding activated hydroxyl amino acid derivatives e.g. mesylate,[18] tosylate,[19,20] chloride,[21] iodide[19] (Denoted as LG in **Scheme 3.8**). Moreover, a variety of thio-nucleophiles (Nu^- in **Scheme 3.8**) can be used depending on the desired function of the target thio-amino acid. Our focus is on obtaining the free amine/thiol hydrochloric acid salt of methyl ester protected amino acid. Therefore, the thio-nucleophile of choice would be thioacetate as the acetate protecting group (**35** or **36**) can be removed under acidic conditions in the same step as *N*-Boc deprotection (*i* in **Scheme 3.8**) to yield the free amine/thiol hydrochloric acid salt of methyl ester protected amino acid.



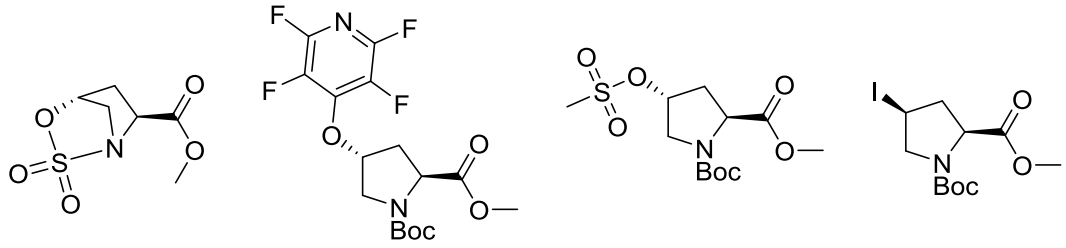
Scheme 3.8: Proposed synthesis of **33** or **34**. *i*: thionyl chloride, methanol, reflux 3 h.

With this in mind we synthesised Boc-Hyp-OMe (**38**) from commercially available (Bachem) (2*S*,4*R*) H-Hyp-OMe.HCl (**37**) starting material (**Scheme 3.9**).



Scheme 3.9: Synthesis of **38**. *i*: di-*tert*-butyl-dicarbonate, MeCN, K_2CO_3, RT, 18 h.

With H-L-Hyp-OMe (**37**) and Boc-L-Hyp-OMe (**38**) in hand we attempted to synthesise a series of activated hydroxyl amino acid derivatives (**39**-**42** in **Figure 3.2**) to investigate different routes to access thioproline.

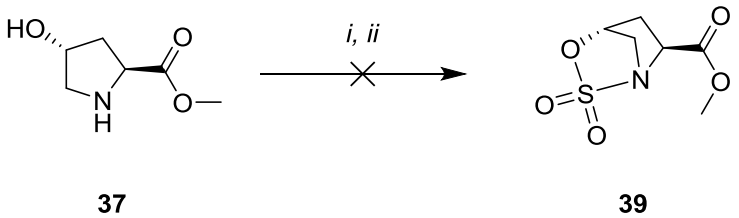


39 **40** **41** **42**

Figure 3.2: Potential intermediate compounds (**39**-**42**) in the syntheses thioproline.

3.2.1.1 Thioproline Syntheses: Cyclic Sulfamidate

Sulfamidate formation and subsequent ring opening with thiol nucleophiles is a highly stereo- and regio-specific strategy that has been shown to work on both serine and threonine amino acids.[22] There are no reports in the literature of preparation of proline sulfamidates. Therefore, the reported reaction conditions[22] were used in an attempt to prepare the bicyclic sulfamidate (**39**) from hydoxy-proline (**37**). All efforts to achieve this transformation failed. Analysis of the crude reaction mixture suggested that the reason for the lack of success was due to an ineffective ring closure step in the initial formation of the cyclic sulfamadite (**Scheme 3.10**).

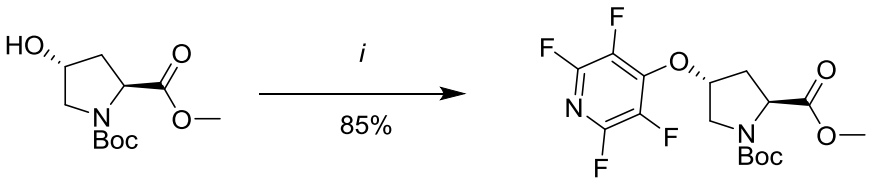


37 **39**

Scheme 3.10: Attempted syntheses of **39**. *i*: Pyridine, DCM, -78 °C; $SOCl_2$, -78 °C →5 °C, 1 h; *ii*: $RuCl_3.3H_2O$, $NaIO_4$, MeCN, 5 °C, 10 min; RT, 10 min.

3.2.1.2 Thioproline Syntheses*:* Tetrafluoropyridine Leaving Group

Previous work within the group has shown that the *trans*-tetrafluoropyridine containing serine amino acid can undergo elimination, substitution, and nucleophilic aromatic substitution reactions depending on the reaction condtions employed. Therefore, we looked to prepare the novel proline compound **40**. **40** was synthesised in a straightforward manner in excellent yield (**Scheme 3.11**) from **38** with retention of stereochemistry. The absolute configuration was solved by X-ray diffraction (**Figure 3.3**) and.



38 **40**

Scheme 3.11: Synthesis of **40**. *i*: pentafluoropyridine, K_2CO_3, MeCN, RT, 48 h.

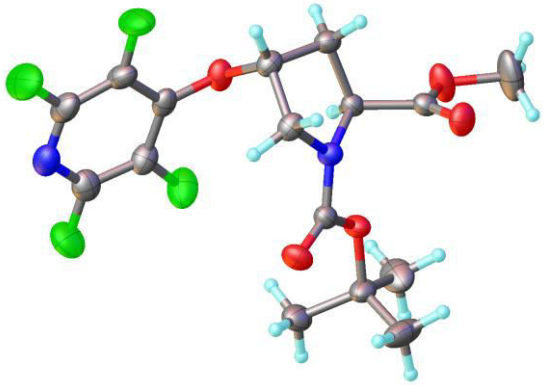
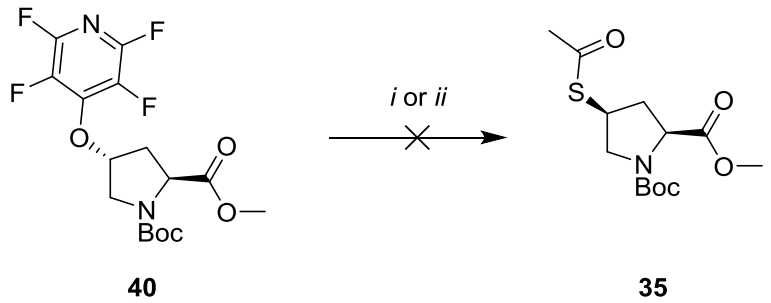


Figure 3.3: Molecular structure of **40** showing 50% probability anisotropic displacement ellipsoids**.**

However, the conditions used in the subsequent thiol nucleophilic substitution (*i* or *ii* in **Scheme 3.12)** with thioacetate yielded no evidence of the desired attack at the γ-carbon. Unreacted starting material as well as small amounts of nucleophilic aromatic substitution were observed *via* $ESI^{+/-}$ LCMS. The reactions of **40** are currently being investigated by another member of the Cobb group.

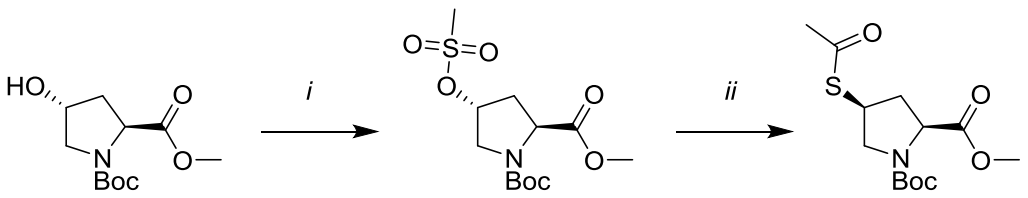


40 **35**

Scheme 3.12: Attempted synthesis of **35.** *i*: potassium thioacetate, DMF, RT, 18 h. *ii*: potassium thioacetate, DMF, 70 °C, 5 h; RT ,18 h.

3.2.1.3 Thioproline Synthesis: Mesylate Leaving Group

Conversion of the hydroxyl group in **38** into a leaving group *via* mesylation followed by nucleophilic attack is a widely used and broadly successful synthetic strategy. Mesylates are known to be unstable, however, *in situ* mesylate generation or use of the crude product by-passes any noticeable degradation and often leads to good yields of the final target compound. This proved to be true as the crude mesylate showed excellent reactivity to potassium thioacetate to the desired S-γC bond in good yield (**Scheme 3.13**).



38 **41** **35**
58% over 2 steps

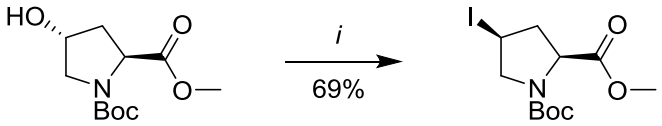
Scheme 3.13: Synthesis of **35**. *i*: methanesulfonyl chloride, Et_3N, DCM, 0 °C, 1 h. *ii*: potassium thioacetate, DMF, 70 °C, 1 h 30 min.

The chemistry outlined in **Scheme 3.13** leads to an inversion of stereochemistry at the γ-carbon, yielding the *cis* proline configuration (**35**). To access the *trans* configuration *via* this route would require the (2*S*,4*S*)-*cis*-L-hydroxproline which can be purchased at £599 per gram (Sigma) or synthesised in 4 steps from *trans*-hydroproline. The published yield over 4 steps is 66% and the resulting *cis*-hydroxyproline would still

need methyl ester protection, mesylation and thiol attack.[23] Given the associated costs an alternative approach was sought.

3.2.1.4 Thioproline Synthesis: Iodine Leaving Group

Literature[24,25] shows that conversion of the hydroxyl to an iodo moiety in an "Appel" type reaction can be controlled to yield either pure *cis* or an isolatable mixture of *cis*- and *trans*-4-iodo-proline depending on the reaction conditions. The literature procedure was followed and Boc-cis-iodoproline-OMe (**42**) was isolated in 69% yield (**Scheme 3.14**) with complete inversion of stereochemistry and no evidence of the trans-isomer being present. The absolute configuration of **42** was solved by X-ray diffraction (**Figure 3.4**).



38 **42**

Scheme 3.14: Synthesis of **42**. *i*: PPh_3, I_2, imidazole, 4 h, RT.

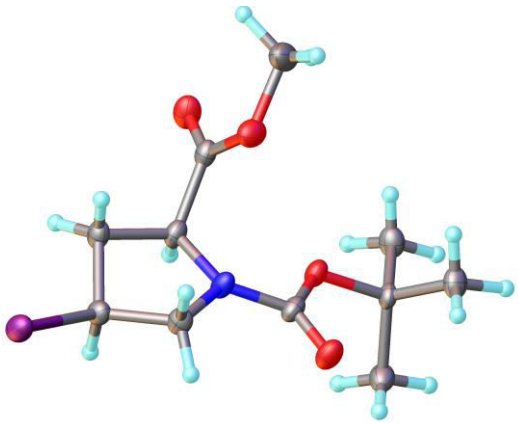
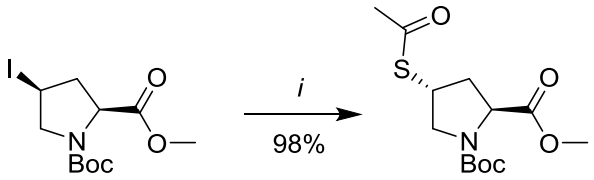


Figure 3.4: Molecular structure of **42** showing 50% probability anisotropic displacement ellipsoids**.**

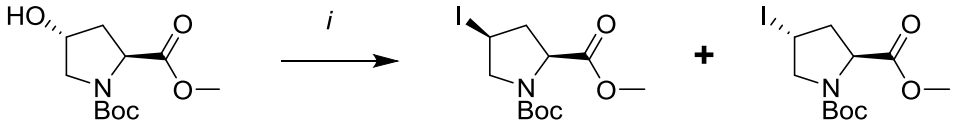
Thiol attack of the *cis*-iodoproline (**42**) resulted in the desired product (**36**) being prepared in an excellent yield (**Scheme 3.15**).



42 **36**

Scheme 3.15: Synthesis of **36**.*i*: potassium thioacetate, DMF, 70 °C, 5 h; RT 18 h.

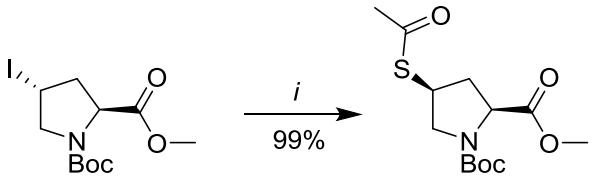
The procedure in **Scheme 3.14** was also carried out under different conditions (increased iodine equivalents) to yield an equilibrium mixture of *cis*-(**42**) and *trans*-Boc-iodoproline-OMe (**43**) (**Scheme 3.16**). The *cis* (**42**) and *trans* (**43**) products were separable by column chromatography and the products were isolated in 38% and 35% yield, respectively.



38 **42** 38% **43** 35%

Scheme 3.16: Syntheses of *cis*- and *trans*-thioprolines: **42** and **43**, respectively. *i*: PPh_3, I_2, imidazole, 48 h, RT.

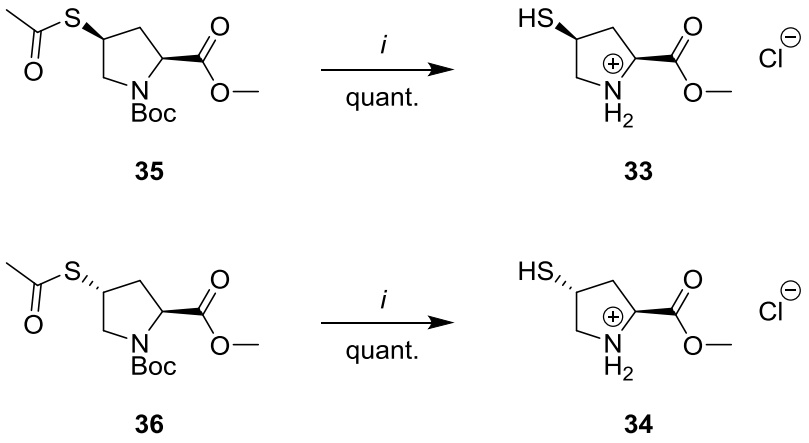
The *trans*-iodo proline was treated with potassium thioacetate yielding **35** in excellent yield (**Scheme 3.17**).



43 **35**

Scheme 3.17: Synthesis of **35**.*i*: potassium thioacetate, DMF, 70 °C, 5 h; RT 18 h.

The compounds **35** and **36** were treated with acidic methanol conditions to yield the deprotected products **33** and **34** in quantitative yields (**Scheme 3.18**).



Scheme 3.18: Synthesis **33** and **34**. *i*: thionyl chloride, methanol, reflux 3 h.

3.2.2 Thiol Amino Acid Synthesis Conclusions

All four compounds in the target library (**31**-**34**) were obtained in yields of over 58% (**Figure 3.5**) from commercial available starting materials. **31** was purchased (Sigma) and **32** was produced in 98% yield in one synthetic step from H-Pen-OH. **33** and **34** were produced in a three step synthesis (mesylation (**33**)/iodoination(**34**), nucleophilc thiol attack and acidic deprotection) from Boc-Hyp-OMe (**38**) in yields of 58% and 68%, respectively. The attempts to synthesize **33** *via* tetrafluoropyridyl (**40**) or cyclic sulfamidate amino acids were unsuccessful. However, the chemistry of **40** is being further investigated. Studies to determine the pK_as of **31** – **34** and thus, potential application in kinetic NCLs were undertaken.

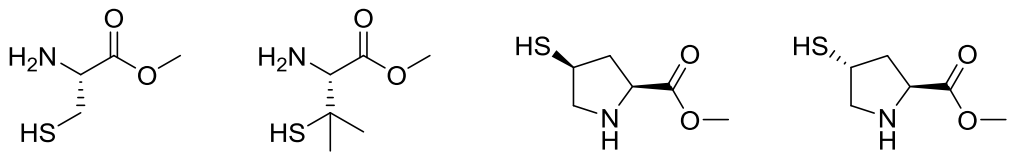


Figure 3.5: Target library of thiol amino acids and their overall synthetic yield from commercial starting materials.

3.3 pK_a Studies on Thiol Containing Amino Acids

3.3.1 pK_a Studies

**The work reported here is a short summary of the investigation carried out by our collaborators Mr O. Maguire a PhD student in the O'Donoghue group within the Chemistry Department at Durham University.*

The four compounds (**31**-**34**) exist in one of four states of protonation depending on their pH environment (**Figure 3.6**). For example, the zwitterion species of **31**, denoted *(-)S*-Cys-OMe-*NH_3*+ in **Figure 3.6** and **Figure 3.7**.

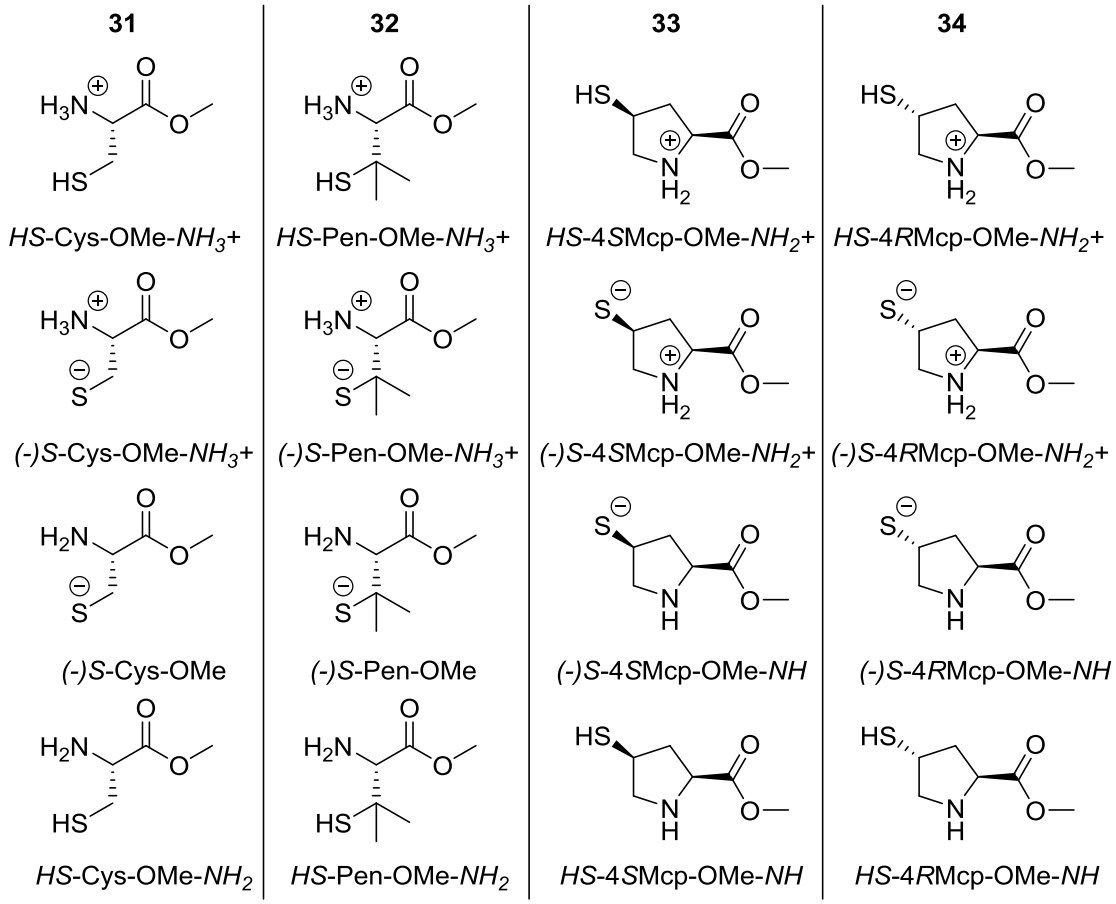


Figure 3.6: Target library of thiol-amino acids and their varying states of protonation.

The absorbance of the four species (of each amino acid) were studied in pH range of 0 – 14 (**Figure 3.7**) and used to calculate dissociation constants and hence, the pK_as of the thiol in the relative species (**Table 3.1**). However, the results for **34** were found to be erroneous (due to thiol oxidation) and are not reported. Further experiments are being undertaken in the presence of TCEP (a reducing agent).

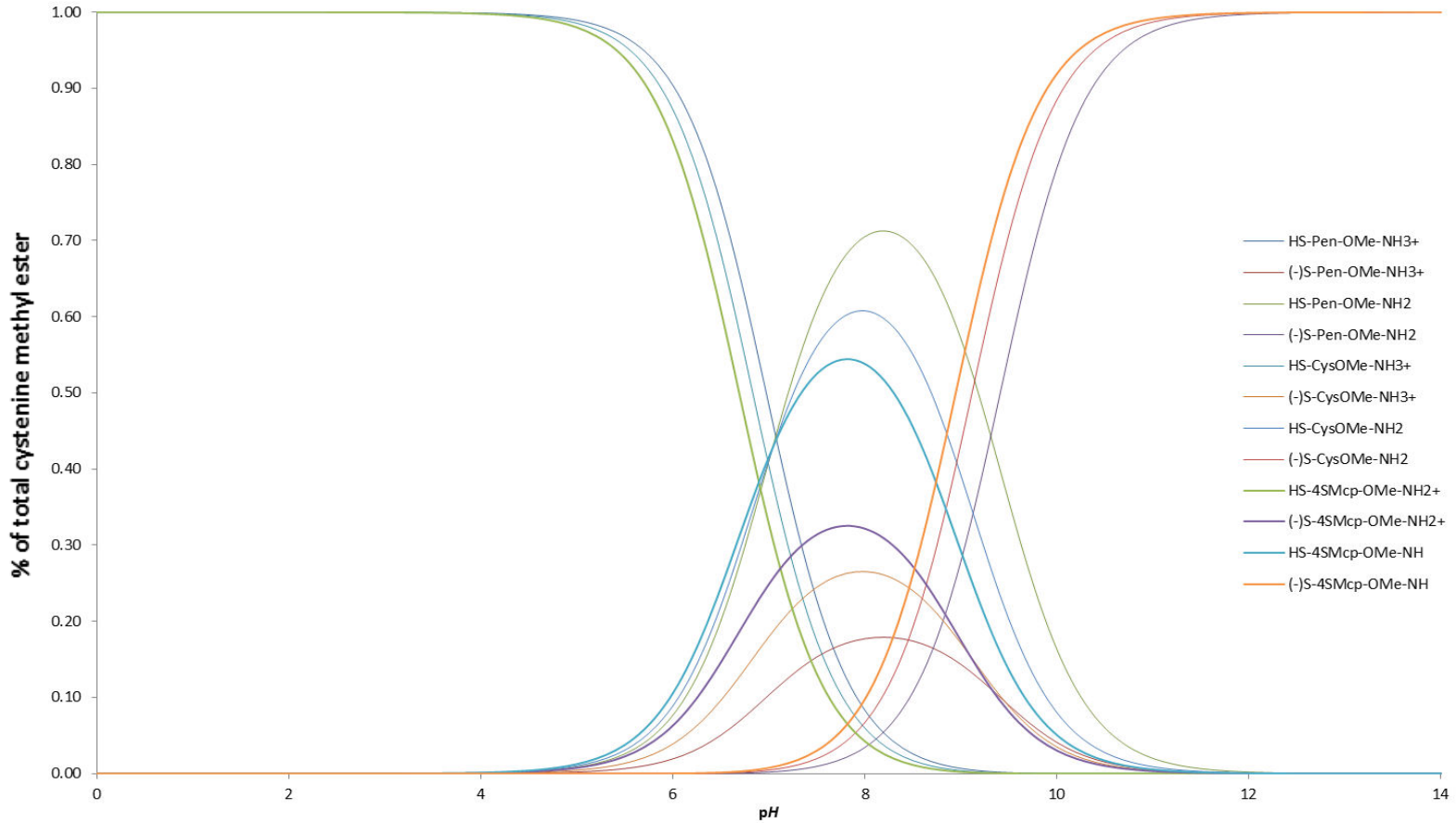


Figure 3.7: The percentage abundance in solution of each of the four species in the pH range 0 – 14 for the cysteine methyl ester (**31**, Cys-OMe), penicillamine methyl ester (**32**, Pen-OMe) and 4*S*-mercaptoproline methyl ester (**33**, 4SMcp-OMe) at 25 °C and ionic strength I = 0.3 M (KCl).

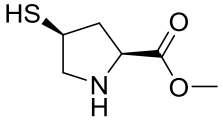
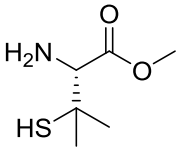
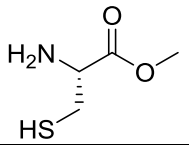
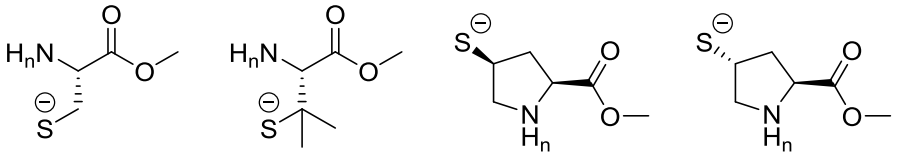
Compound Number	Structure	pK_A H_3N^+-R-SH ⇌ H_3N^+-R-S^-	pK_B H_3N^+-R-SH ⇌ H_2N-R-SH	pK_C H_3N^+-R-S^- ⇌ H_2N-R-S^-	pK_D H_2N-R-SH ⇌ H_2N-R-S^-
31		7.35	6.99	8.60	8.95
32		7.67	7.07	8.71	9.31
33		7.12	6.89	8.52	8.74

Table 3.1: pK_a estimates for the amino acids **31**, **32** and **33** at 25 °C and ionic strength I = 0.3 M (KCl).

3.3.2 pK_a Studies: Conclusions

The studies show significant differences in the pK_a values of **31**, **32** and **33** and hence, abundance of ionic species at a given pH. In relation to NCL reaction the abundance of thiolate species (either zwitterionic or anionic) is of interest (**Figure 3.8**).



	31	**32**	**33**	**34***
pH 6.8	15%	8%	21%	-
pH 7.2	22%	13%	24%	-

Figure 3.8: The abundance of thiolate species (either zwitterionic or anionic) present at pH 6.8 and 7.2.***34*** provided erroneous results due to oxidation.

The thiolate species is thought to be responsible for initial nucleophilic attack upon the thioester in an NCL reaction. Thus, a low abundance of thiolate at a particular pH

could render the NCL of one amino acid negligible and whilst another (with higher abundance) may proceed. The correlated pK_a data which leads to calculated abundances of thiolate species could therefore lead to facile planning and execution of kinetic "one-pot" ligations with these amino acids.

3.5 References

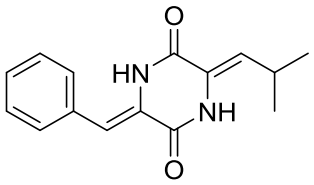
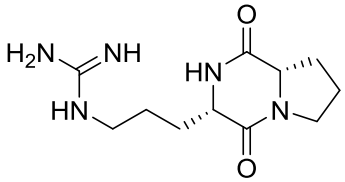
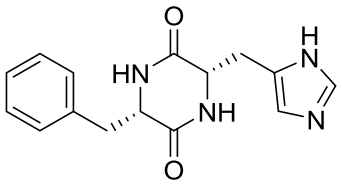
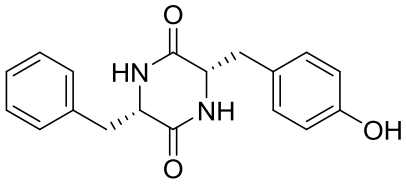
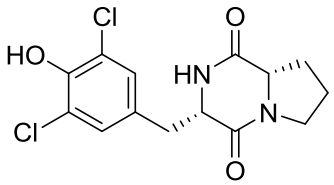
1. P. Dawson, T. Muir, I. Clark-Lewis and S. Kent, *Science*, 1994, **266**, 776-779.
2. T. W. Muir, *Ann. Rev. Biochem.*, 2003, **72**, 249-289.
3. J. P. Richardson and D. Macmillan, *Semi-synthesis of Glycoproteins from E. coli Through Native Chemical Ligation*, Humana Press Inc, Totowa, 2011.
4. T. Wieland, E. Bokelmann, L. Bauer, H. U. Lang and H. Lau, *Justus Liebigs Ann. Chem.*, 1953, **583**, 129-149.
5. S. B. H. Kent, *Chem. Soc. Rev.*, 2009, **38**, 338-351.
6. C. Haase, H. Rohde and O. Seitz, *Angew. Chem. Int. Ed.*, 2008, **47**, 6807-6810.
7. L. Z. Yan and P. E. Dawson, *J. Am. Chem. Soc.*, 2001, **123**, 526-533.
8. B. L. Pentelute and S. B. Kent, *Org. Lett.*, 2007, **9**, 687-690.
9. Q. Wan and S. J. Danishefsky, *Angew. Chem. Int. Ed.*, 2007, **46**, 9248-9252.
10. Z. Harpaz, P. Siman, K. S. Kumar and A. Brik, *Chembiochem*, 2010, **11**, 1232-1235.
11. S. D. Townsend, Z. P. Tan, S. W. Dong, S. Y. Shang, J. A. Brailsford and S. J. Danishefsky, *J. Am. Chem. Soc.*, 2012, **134**, 3912-3916.
12. D. Crich and A. Banerjee, *J. Am. Chem. Soc.*, 2007, **129**, 10064-+.
13. D. Bang and S. B. H. Kent, *Angew. Chem. Int. Ed.*, 2004, **43**, 2534-2538.
14. D. Bang, B. L. Pentelute and S. B. H. Kent, *Angew. Chem. Int. Ed.*, 2006, **45**, 3985-3988.
15. Z. Tan, S. Shang and S. J. Danishefsky, *Angew. Chem. Int. Ed.*, 2010, **49**, 9500-9503.
16. J. Chen, P. Wang, J. Zhu, Q. Wan and S. J. Danishefsky, *Tetrahedron*, 2010, **66**, 2277-2283.
17. R. Yang, K. K. Pasunooti, F. Li, X. W. Liu and C. F. Liu, *J. Am. Chem. Soc.*, 2009, **131**, 13592-13593.
18. L. Isakovic, O. M. Saavedra, D. B. Llewellyn, S. Claridge, L. Zhan, N. Bernstein, A. Vaisburg, N. Elowe, A. J. Petschner, J. Rahil, N. Beaulieu, F. Gauthier, A. R. MacLeod, D. Delorme, J. M. Besterman and A. Wahhab, *Bioorg. Med. Chem. Lett.*, 2009, **19**, 2742-2746.
19. R. Caputo, M. DellaGreca, I. de Paola, D. Mastroianni and L. Longobardo, *Amino Acids*, 2010, **38**, 305-310.
20. V. Eswarakrishnan and L. Field, *J. Org. Chem.*, 1981, **46**, 4182-4187.
21. Y. Berger, H. Dehmlow, D. Blum-Kaelin, E. A. Kitas, B.-M. Löffler, J. D. Aebi and L. Juillerat-Jeanneret, *J. Med. Chem.*, 2004, **48**, 483-498.
22. S. L. Cobb and J. C. Vederas, *Org. Biomol. Chem.*, 2007, **5**, 1031-1038.
23. M. S. Chorghade, D. K. Mohapatra, G. Sahoo, M. K. Gurjar, M. V. Mandlecha, N. Bhoite, S. Moghe and R. T. Raines, *J. Fluorine Chem.*, 2008, **129**, 781-784.
24. J. R. Dormoy, *Synthesis*, 1982, **9**, 753-756.
25. J.-R. Dormoy and B. Castro, *Synthesis*, 1986, **1**, 81-82.

Chapter 4 : Diketopiperazines - Synthesis and Inhibitory Activity Against CCL2 Induced Chemotaxis.

4.1 An Introduction to 2,5-Diketopiperazines

4.1.1 2,5-Diketopiperazines: A Structural Overview

2,5-Diketopiperazines (DKPs) are cyclic dipeptides that are biosynthesised by a variety of organisms (including bacteria, fungi and mammals) as secondary metabolites.[1] Research interest in DKPs has recently increased due to the wide-ranging bioactivity that many natural occurring systems exhibit (**Table 4.1**):[2,3] e.g. antibacterial,[4-7] antitumour,[7] antifungal[7,8] and antiviral.[9]

Entry	Name	Structure	Biological Target	Refs
1	cyclo(ΔPhe-ΔLeu) (Albonoursin)		antibacterial, antitumor (cell division inhibition)	4,5
2	cyclo(L-Arg-L-Pro)		antifungal (chitinase inhibition)	8
3	cyclo(L-Phe-L-His)		heart rate/coronary flow, antifungal, antibacterial and antitumor	7
4	cyclo(L-Phe-L-Tyr)		μ-opiod binding	10
5	cyclo(13,15-dichloro-L-Tyr-L-Pro)		CCL2 induced chemotaxis inhibition	11

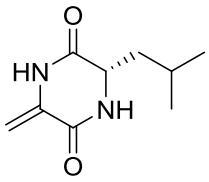
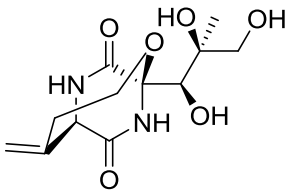
6	cyclo(DHA-L-Leu)		antihyperglycaemic (yeast and porcine α-glucosidase inhibition)	12
7	Bicyclomycin		antibacterial (inhibition of transcription termination factor - Rho)	6

Table 4.1: The broad spectrum bioactivity of naturally occurring DKPs.

DKPs contain two *cis*-amide bonds (hydrogen bond donors/acceptors) and with the exception of glycine containing DKPs, two stereo centres (chiral α-carbons,* in **Figure 4.1).** The chirality combined with the rigid amide bonds specifically orientates the amino acid side chains in space around a cyclic core (R_1 and R_2 in **Figure 4.1**).

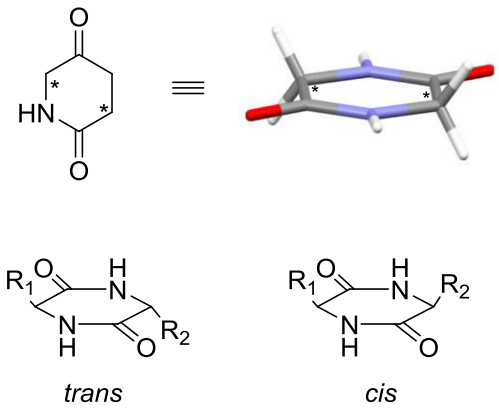


Figure 4.1: The structure of the 2,5-diketopiperazine: cyclo(Gly-Gly) with α-carbons indicted with *.[13] When chiral amino acid residues are present the stereochemistry defines either a *trans*- and *cis*- arrangement of R_1/R_2 side-chains.

DKPs generally have a planar 6-membered ring conformation[14] or a slightly distorted boat conformation with a small energy barrier (1-5 kcal/mol) between the forms.[2] For example, cyclo(L-Asp-L-Asp) forms a boat conformation with pseudo C2v symmetry[15] and although cyclo(Gly-Gly) shows a planar crystalline structure[16] gas phase studies have indicated a preference for a boat conformation.[17] The core 6-membered ring is often distorted by specific substituents, especially aromatic and proline residues. The interaction between the polarizable π-electron cloud of an aromatic ring and the amide

bond dipoles lead to a small preference for a folded structure where the substituent overlaps the DKP core (e.g. cyclo(Gly-L-Phe) **Figure 4.2**).[18] In the case of proline containing DKPs the cyclic nature of the amino acid imparts significant ring constraints that favour a boat conformation (e.g. cyclo(Gly-L-Pro) **Figure 4.2**).[2,19] The natural product CCL2 induced chemotaxis inhibitor: cyclo(13,15-dichloro-L-Tyr-L-Pro) (*entry 5*, **Table 4.1**) contains both aromatic and proline residues, yet the conformation is not known.

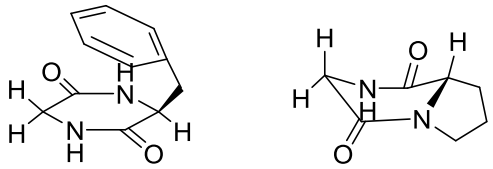


Figure 4.2: folded and boat structures of cyclo(Gly-L-Phe) and cyclo(Gly-L-Pro). ^{1}H NMR indicates that aromatic DKPs tend to form a folded structure[18] and computational conformational mapping shows that bicyclic proline containing DKPs are restricted to a small area of conformational space and a boat conformation.

The conformational aspects combined with the variety of possible side chain moieties give DKPs a broad bioactivity and therapeutic spectrum. This has culminated in the successful development of several clinically approved drug molecules containing the DKP core scaffold: Epelsiban[20] and Plinabulin in **Figure 4.3**.[21] Epelsiban is a selective oxytocin receptor antagonist and Plinabulin is a vascular disrupting agent that has completed phase I and phase II clinical trials for non-small cell lung cancer.[22-24] Plinabulin is an achiral analogue of the DKP metabolite: Phenylhistatin (isolated from *Aspergillus ustus*).[25]

For a comprehensive recent review of the area see: "Diketopiperazines: Biological Activity and Synthesis"[1] or "2,5-Diketopiperazines: Synthesis, Reactions, Medicinal Chemistry, and Bioactive Natural Products".[2]

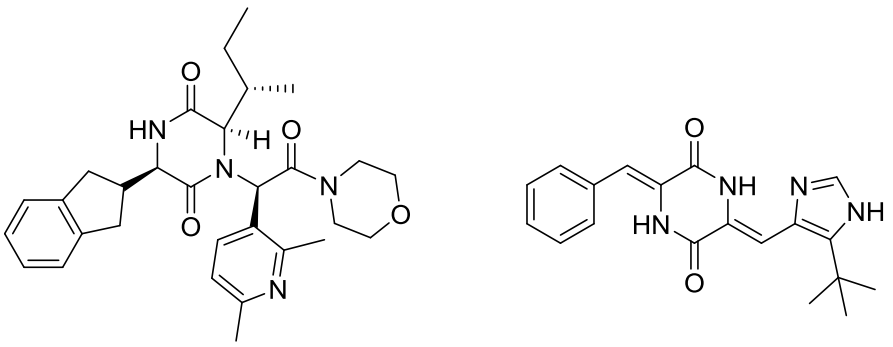
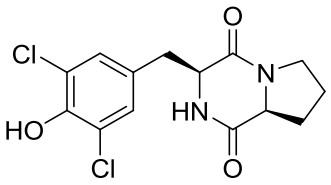


Figure 4.3: Structure of Epelsiban and Plinabulin.

4.1.2 DKPs as Inhibitors of Chemokine Mediated Chemotaxis

Chemokines are small proteins that play a huge role in a number of allergic, autoimmune, inflammatory, and viral diseases and in America 50 million people suffer from allergic diseases per year (for which treatment costs around $18 billion).[26] The study of small-molecule chemokine receptor antagonists as anti-inflammatory agents is a growing field that is still very much in its infancy. Hence, a limited number of studies to evaluate potential efficacy of chemokine antagonists have been reported.[27,28] However, it took only 11 years from the discovery of the chemokine receptor: CCR5 to the FDA approval of small-molecule inhibitor: Maraviroc. A year later Plerixafor was FDA approved as a CXCR4 antagonist for hematopoietic stem cell mobilization.

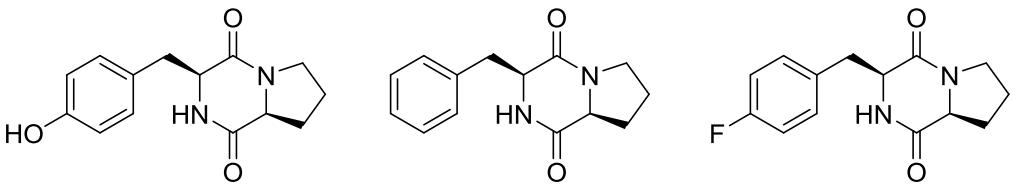
Previously published[29] work from the Cobb group confirmed the ability of naturally occurring DKP: cyclo(13,15-dichloro-L-Pro-L-Tyr) (**46**, **Figure 4.4**) to act as a chemokine receptor antagonist and inhibit CCL2-induced chemotaxis *in vitro*.[11]



46
cyclo(13,15-dichloro-L-Pro-L-Tyr)

Figure 4.4: Structure of cyclo(13,15-dichloro-L-Pro-L-Tyr) **46**. **46** is isolated from the fungus *Leptoxyphium* sp.

Based on **46**, a library of 8 DKPs were successfully synthesised and tested in chemotaxis assays (**Chapter 1**, **Section 1.4.3** and **1.4.4**). The work indicated that the substituents on the aromatic ring are important for activity: The activity of the non-chlorinated analogue: cyclo(L-Pro-L-Tyr) (**47, Figure 4.5**) was found to be significantly reduced, whereas cyclo(L-Phe-L-Pro) (**48, Figure 4.5**) and cyclo(*p*-fluoro-L-Phe-L-Pro) (**49, Figure 4.5**) showed similar levels of inhibition as the natural product **46** (**Figure 4.6**).



47
cyclo(L-Pro-L-Tyr)

48
cyclo(L-Phe-L-Pro)

49
cyclo(*p*-fluoro-L-Phe-L-Pro)

Figure 4.5: Structure of cyclo(L-Pro-L-Tyr) **47**, cyclo(L-Phe-L-Pro) 48 and cyclo(*p*-fluoro-L-Phe-L-Pro) **49**.

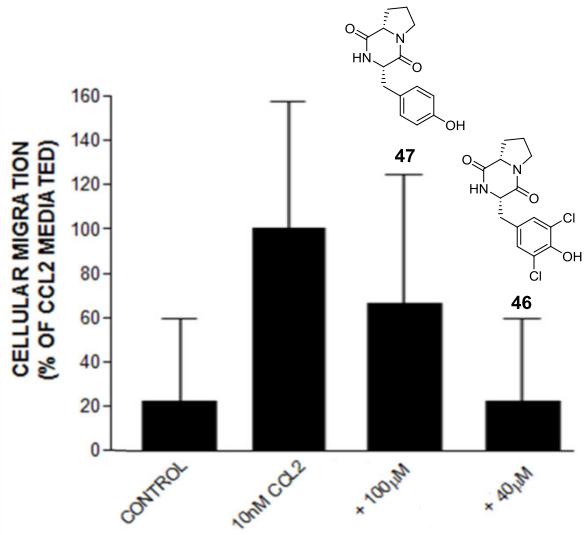


Figure 4.6: Cellular migration of THP-1 cells in the prescence 10 nM CCL2 and DKPs **46** and **47** at concentration of 40 and 100 μM, respectively. The control contains no CCL2 or DKP. The data is normalised to 100% with 10nM CCL2 in the absence of inhibitor.

Multiple attempts were made to crystallise all compounds in the library so that biophysical studies (including modelling/docking) could be undertaken and future inhibitors could be intelligently designed. However, crystals of a good enough quality to solve were formed by three DKPs (**47**, **48** and **49**, **Figure 4.7**) with the natural product **46** and five other DKPs remaining undefined. Although as yet modelling/docking with CCL2 have not been attempted, the crystal structure of the non-active DKP **47** shows the typical folded structure of aromatic DKPs whereas the active* DKPs **48** and **49** have elongated structures where the only significant variation is in the angle that the aromatic ring is orientated compared to the bicycle.

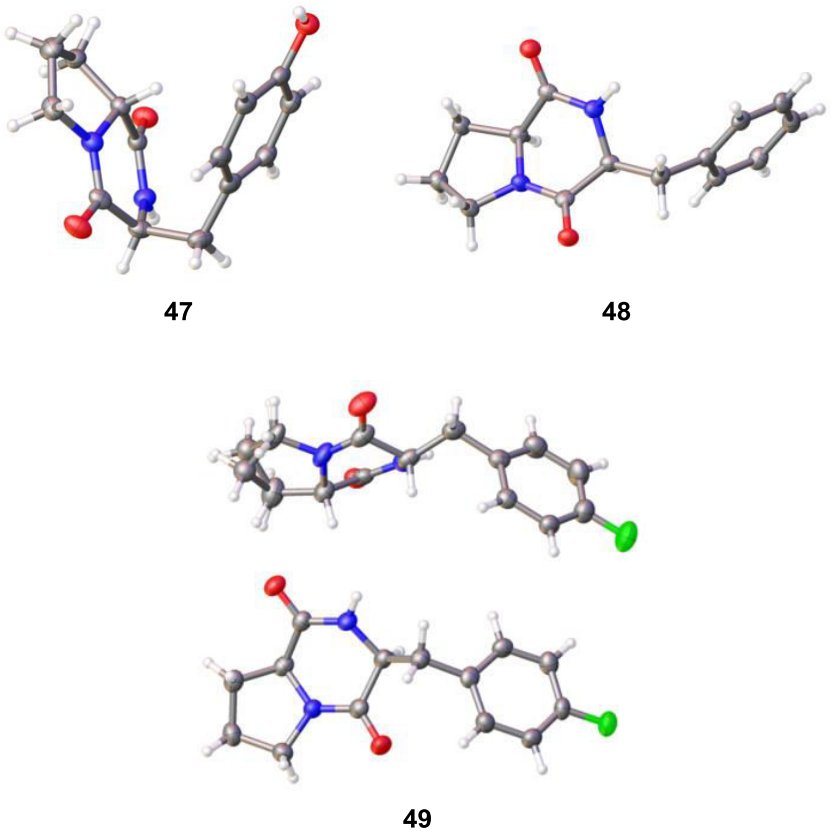


Figure 4.7: Molecular structures of **47**, **48** and **49** showing 50% probability anisotropic displacement ellipsoids**.** The unit cell in the crystal structure of **49** contains two molecules and shows slight disorder in the proline ring.

*__48__ and __49__ were found to give significant inhibition of CCL2 induced chemotaxis at 100 μM, (reduced cellular migration by <40%) whereas __47__ gives no significant inhibition.

Selectivity studies were also undertaken and **49** was found to selectively inhibit CCL2 induced chemotaxis at a concentration of 50 μM but have no effect on CCL7 and CCL5 ligands, which are also known to bind CCR2 and initiate chemotaxis (**Figure 4.8**).

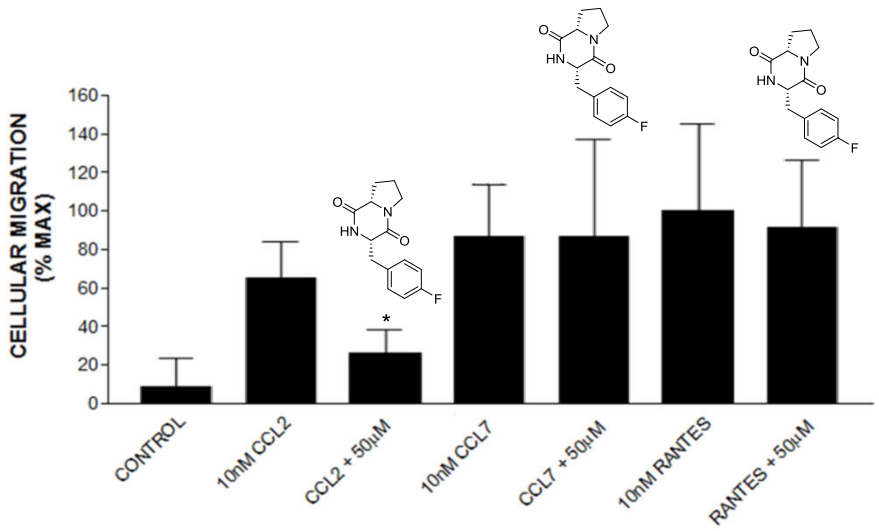


Figure 4.8: Cellular migration of THP-1 cells in the prescence 10 nM CCL2, CCL7 and CCL5 (RANTES) and at concentration of 50 μM. The control contains no chemokine or DKP. The data is normalised to 100% with 10nM CCL5 in the absence of inhibitor.

These initial results strongly suggest that the DKPs prepared do not disrupt the CCL2-CCR2 protein-protein interaction and that the mechanism of inhibition does not involve binding to either the receptor (CCR2) or the receptor binding site on the ligand (CCL2, CCL7 and CCL5). Therefore, It is believed that the DKPs inhibit CCL2 chemotaxis by binding specifically to CCL2 and at non-receptor binding site. For example, the inhibition could originate *via* binding at the dimerization face or a site that causes an allosteric effect at the receptor binding site

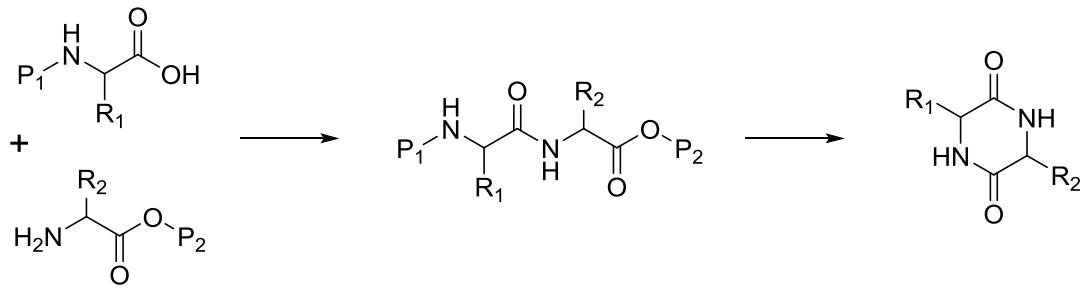
Although the acitivity of DKPs would need to be enhanced for theraupeutic use, the interesting CCL2 specificity could grant these molecules application as chemical probes. In particular, such probes would be useful as a means of shutting down CCL2 action from the network of chemokine interactions that increase the complexitity of studying aspects of specific chemokine function.[29] The limited number of compounds in the DKP library studied mean that no clear structure activity relationship could be determined. Hence, there is a need to further expand the library of analogs to define the

structural features required for activity and to provide more insights into the mode of CCL2 induced chemotaxis inhibition.

4.2 Introduction to the Synthesis of DKPs

The two most prevalent strategies for the chemical synthesis of 2,5-DKPS are the dipeptide ester cyclisation (**Scheme 4.1**) and the a 4-component Ugi reaction (**Scheme 4.1**). However, other less prevalent strategies exist such as direct amino acid condensation[30,31] or aza-wittig cyclisation.[32]

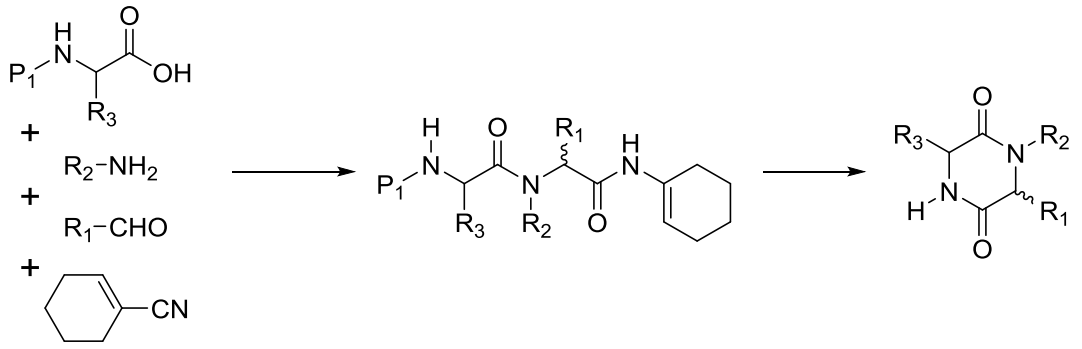
In the dipeptide ester strategy an orthogonally protected dipeptide is synthesised using peptide coupling reagents then subsequently cyclised. The cyclisation step requires long reaction times, heating in acid, base or high boiling point solvents which can in some cases can cause racemisation.[33-36] However, advances in microwave assisted heating[36] and solid supported synthesis[37] have provided high yielding DKPs without racemisation.



Scheme 4.1: The general scheme for dipeptide ester cyclisation leading to DKP preparation.

The Ugi reaction does not require peptide coupling reagents yet the terminal amide formed leads to difficult cyclisation steps.[38] Post condensational modification (PCM) Ugi reactions have overcome this and provide fast and efficient DKP synthesis in solution or on a solid support. Routinely an isonitrile precursor is used to form an activated amide that under acidic postcondensational conditions converts to a transient *N*-acylimminium ion. [38-41] This strategy is also applicable to solid supported synthesis.[42] Recently the Ugi method was used superbly in the one-pot synthesis of the natural

product: Thaxtomin A. The only drawback is the lack of stereocontrol inherent to the Ugi reaction.[43]

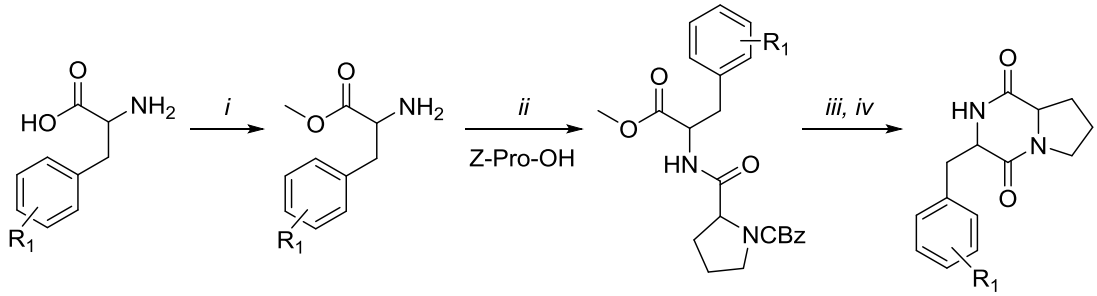


Scheme 4.2: The general scheme for a 4-component PCM Ugi synthesis of a DKP.

For an expansive review in this area see: "2,5-Diketopiperazines: Synthesis, Reactions, Medicinal Chemistry, and Bioactive Natural Products".[2] This review outlines simple synthetic strategies that have led to vast libraries of simple DKPs (including *entries 1*, *2*, *3*, *4* in **Table 4.1**) as well as detailed synthesis of complex DKP derivatives e.g. gliotoxins, aspirochlorines and austamides.

4.2.1 Solution Phase Dipeptide Ester Cyclisation Synthesis of DKPs

Within the Cobb group a dipeptide ester synthesis strategy was first used for the synthesis of a library aromatic derivatives based on the natural product **46**.[29] This general strategy is prevalent in literature and responsible for the majority of the natural product 2,5-DKP syntheses.[2] The major steps are outlined below (**Scheme 4.3**).



Scheme 4.3: *i*: Solution phase synthesis of a DKP *via* a dipeptide ester. *i*: Acetyl chloride, MeOH, 5 h, reflux (>80%). *ii*: DIC, DIPEA, DCM, RT, 20 h (55 - 80%). *iii*: NH_4HCO_2, Pd, MeOH, 5 h, reflux (>95%). *iv*: MeOH, 48 - 96 h, reflux (>70%).

The synthesis enabled the production of a library of eight DKPs including the aforementioned DKPs **46**, **47**, **48** and **49** (**Figure 4.9**).

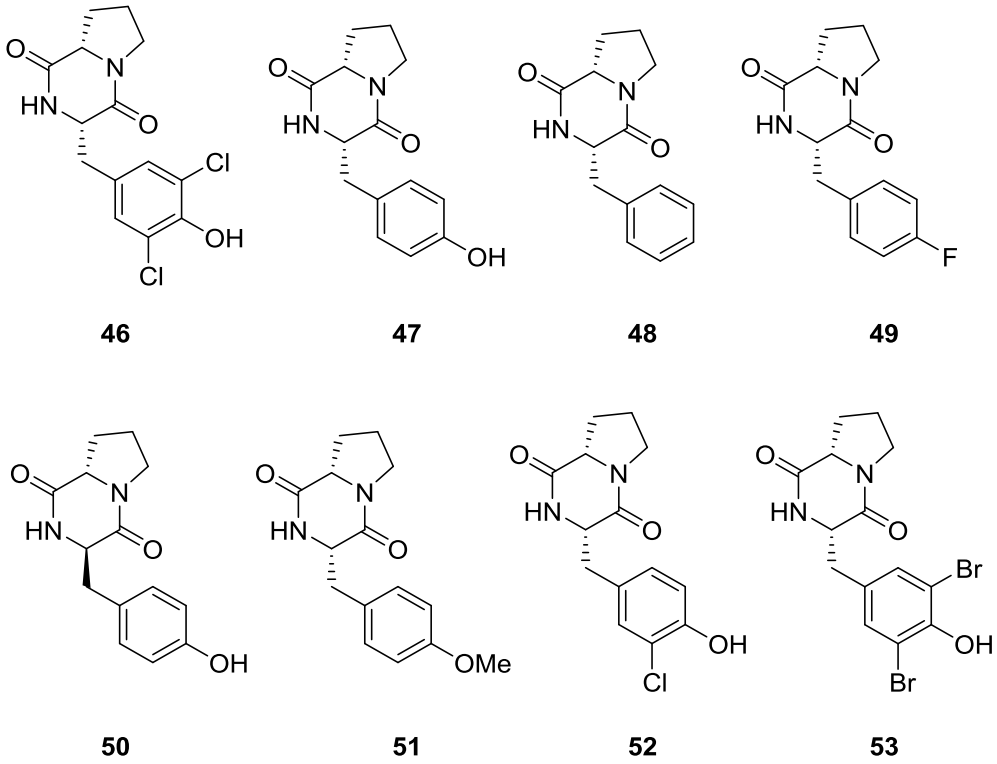


Figure 4.9: The library of eight DKPs synthesised *via* a dipeptide ester solution phase strategy.

Cyclo(L-Pro-D-Tyr) **50** was not a desired product and was formed as an isolatable impurity in the synthesis of **47**. Therefore, notable racemisation occurred during the cyclisation step in this specific case.

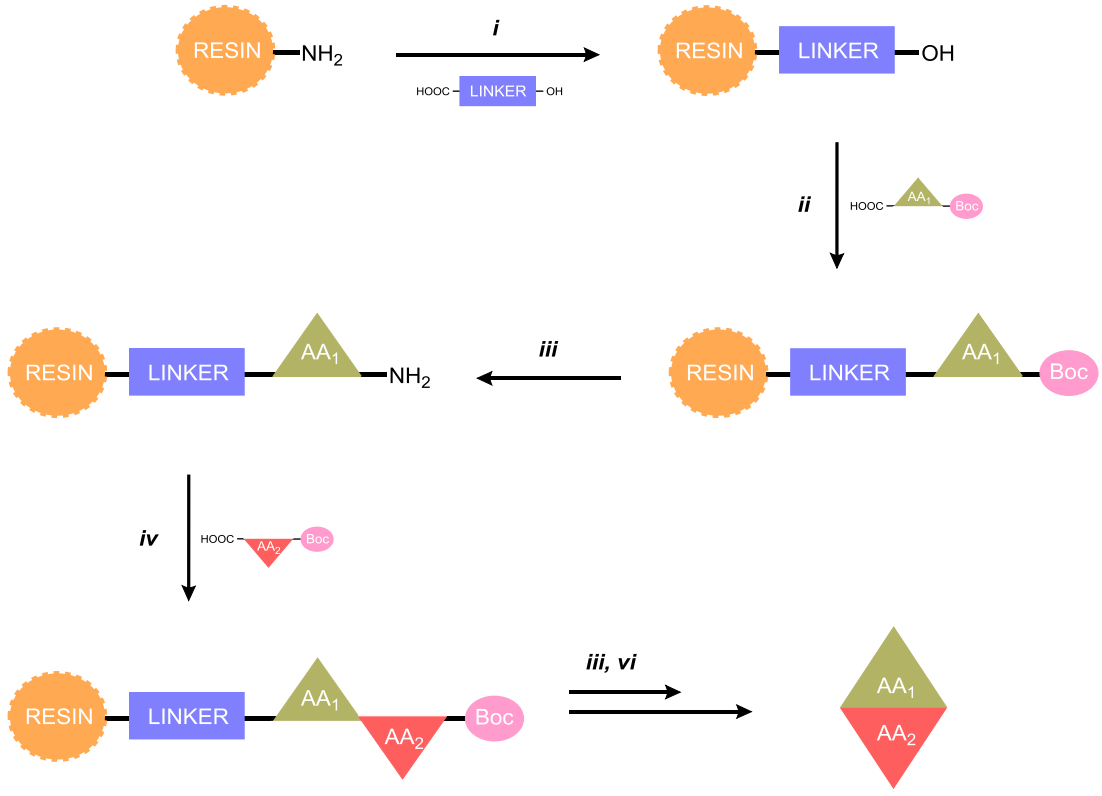
To expand this work and explore the results a new library of DKPs with a wider scope of moieties must be synthesised, tested and crystallised. However, the aforementioned synthesis is time consuming, required elevated temperatures, afforded incomplete stereocontrol and two purification steps (after *ii* and *iv*, **Scheme 4.3**). Therefore, an improved route was devised in an attempt to overcome these problems.

The key advances in this field are the use of microwave heating and solid supports.[44] Thus, a solid supported (on-resin cyclisation) strategy was chosen where microwave

assisted heating could be used in the event of an inefficient step. The strategy was based on a successful on-resin cyclisation and release synthesis used by the Giralt group.[45]

4.2.3 On-Resin Cyclisation and Release Synthesis of DKPs

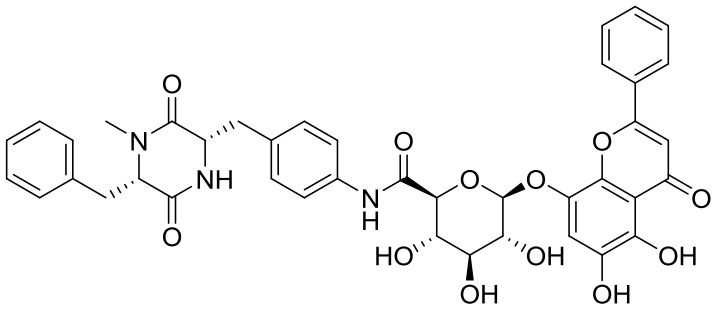
This strategy utilises a reactive linker that once attached to a solid support(resin) allows intramolecular attack in a cleavage/cyclisation step. The general method firstly attaches the linker to the solid support (resin) (*i* in **Scheme 4.4**) and then subsequently forms a dipeptide with *N*-Boc based SPPS chemistry (*ii*, *ii* and *iv* in **Scheme 4.4**). This resin-bound dipeptide is then deprotected to form a protonated terminal amine (*iii* in **Scheme 4.4**). When the terminal amine is neutralised it performs a nucleophilic attack onto the linker moiety. This intramolecular nucleophilic attack leads to the formation of the cyclic DKP (*vi* in **Scheme 4.4**).



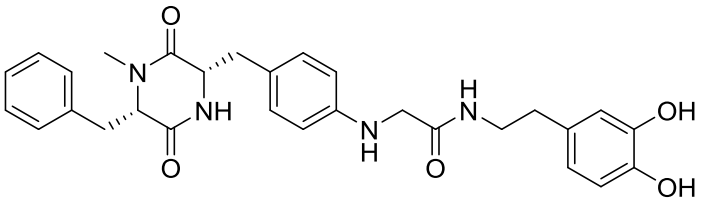
Scheme 4.4: General on-resin synthesis of a Cyclo (AA_1-AA_2) DKP.

Using this strategy the Giralt group synthesised 15 *N*-methylated DKPs and extended the method to the synthesis of complex DKP based biomolecules (e.g. DKP-Dopamine

and DKP-Baicalin) (**Figure 4.10**) as potential blood-brain barrier (BBB) transport shuttles. In the majority of cases the crude DKP obtained after cyclisation and release was deemed pure enough for BBB analysis. However, the reported yields are low (5 - 30%).[45]



DKP-Baicalin



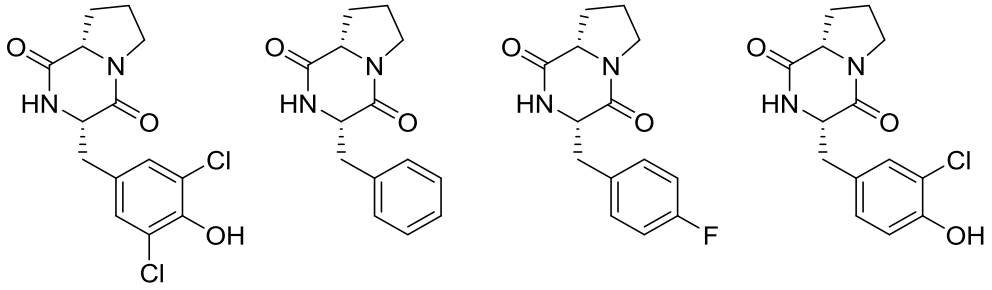
DKP-Dopamine

Figure 4.10: complex DKP based biomolecules synthesised by Geralt *et. al.*: DKP-N-MePhe-Dopamine and DKP-N-MePhe-Baicalin.

4.3 Preparation of a New Library of DKPs

**Aspects of this work were carried out by two undergraduate students under my supervision: Mr Andrew Steer and Ms. Malogrzata Przeradzka.*

The previously synthesised DKP library (**Figure 4.9**) was specifically designed to investigate the effect of varying aromatic ring substituents on the efficacy of the DKPs in CCL2 induced chemotaxis. Hence, all compounds were structurally very similar to the natural product (**46**) scaffold and not surprisingly a number of compounds had comparable activity (**Figure 4.11**). Perhaps the most important result is the high activity of cyclo(L-Pro-L-Phe) (**48**) which is lacking in any phenyl ring substituents. This result initiated interest in performing experiments to further explore the structural limits required to retain chemotactic activity.



46 **48** **49** **52**

Figure 4.11: DKPs that showed significant inhibition of CCL2 induced chemotaxis in a BMC assay at a concentration of 100 μM or lower.

To expand on this previous work a new library was proposed (**Figure 4.12**), specific DKPs moieties (or lack of) were chosen to help further define the important structural features that lead to inhibitory activity. The structural features that we will investigate are stereochemistry (**54**, **55**, **56**), *N*-methylation (**57**), proline substitution (**58**, **59**, **60**) and aromatic substitution (**61**, **62**, **63**).

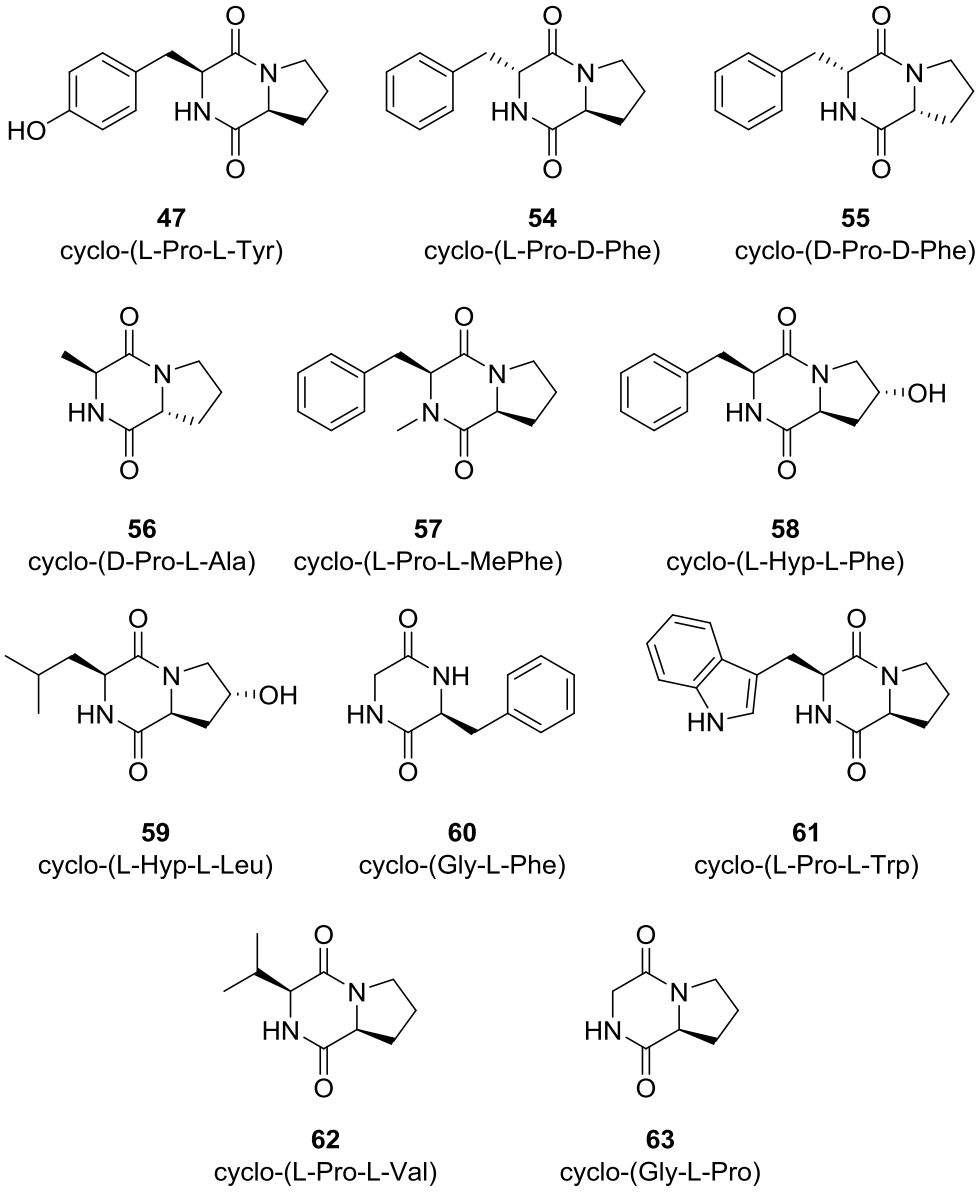


Figure 4.12: The proposed library of DKPs to be tested against CCL2 induced chemotaxis

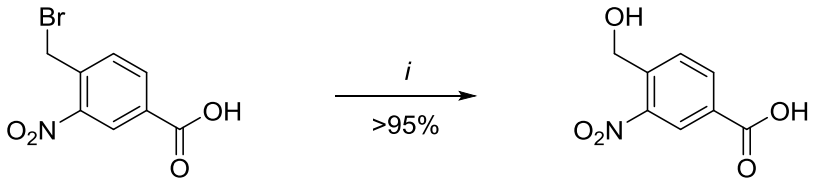
Recently, some simple DKPs have become commercially available (Bachem) including two members of this library: Cyclo(Gly-L-Phe) (**60**) and cyclo(Gly-L-Pro) (**63**). **60** and **63** will show the effect of complete removal of either the proline or aromatic side-chain, respectively. Therefore, they were purchased and tested.

Cyclo(L-Tyr-L-Pro) (**47**) was included in the library, yet it was previously synthesised and proved to be a poor inhibitor when tested. The previous synthesis of this DKP did not control stereochemistry of the tyrosine α-carbon. Therefore, synthesis of this DKP will give an indication of the amount of stereocontrol in the syntheses as well as enabling the DKP to be used as a negative control in the assay.

4.4 On-Resin Synthesis of 2nd Generation of DKPs

4.4.1 Resin Loading and Synthesis of the Linker

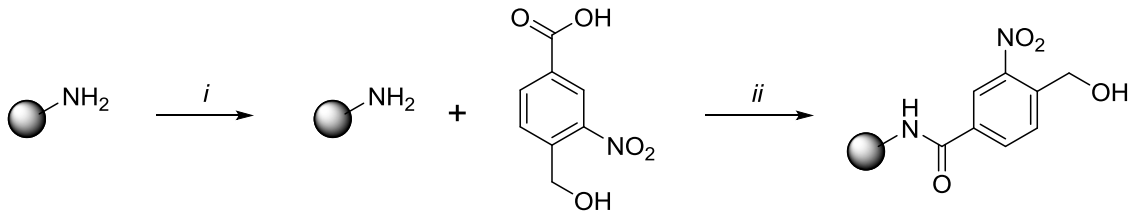
Following the method outlined by Geralt *et al,*.[45] the first step is the preparation of a suitable solid support to facilitate the cyclisation final step. This involves the synthesis of a linker (**64, Scheme 4.5**) and the subsequent loading of **64** onto 4-Methylbenzhydrylamine (MHBA) resin (**Scheme 4.7**) to form MHBA-linker (**65**). **64** was synthesised from 4-bromomethyl-3-nitrobenzoic acid *via* a basic aqueous reflux in near quantitive yield (**Scheme 4.5**).



64

Scheme 4.5: Synthesis of the linker: **65**. *i*: $NaHCO_3$(aq), 30 min, reflux.

The linker (**64**) carboxylic acid was coupled to an extensively washed (*i*, **Scheme 4.7**) MHBA resin using a peptide coupling reagent (DIC). This reaction formed an amide linked pre-loaded resin: MHBA-linker (**65**).



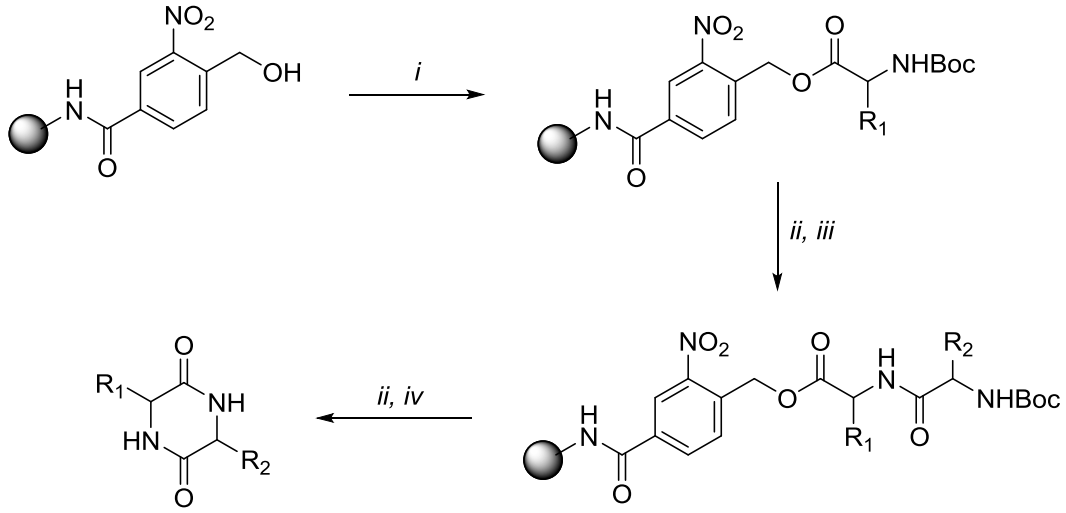
64 **65**

Scheme 4.6: Synthesis of the MHBA- linker: **65**. *i*: DCM, RT, x 4; TFA:DCM (40% v/v), 1 min, RT; TFA:DCM (40% v/v), 5 min, RT, x 2; DIPEA:DCM (5% v/v), 2 min x 3; DCM, RT, x 3. *ii*: DIC, DIPEA, DCM, 18 h, RT.

The activated resin (**65**) was synthesised in gram quantities to enable a number of DKPs to be synthesised from commercially available *N*-Boc amino acids.

4.4.2 On-Resin Synthesis of The DKP Library

The MHBA-linker (**65**) was used to synthesise the library DKPs (**Scheme 4.7**). Our aim was to produce libraries of compounds quickly, economically and in a non-labour intensive way. We were able to optimise and simplify the procedure used by Giralt: excesses of reagents were halved, colorimetric resin tests removed and triple couplings were replaced with double couplings. Additionally, the optimisations had no noticeable effect on the yield of pure DKP (**Table 4.2**) and enabled the synthesis, purification and analysis of a particular DKP to be completed within 8 hours.



Scheme 4.7: Solid phase synthesis of a DKP *via* the cyclisation and release strategy. *i*: Boc-AA(R_1)-OH, DIC, DMAP, DCM, 30 min, RT, x 2. *ii*: TFA:DCM (40% v/v), 10 min, RT, x 2. *iii*: Boc-AA(R_2)-OH, PyBOP, DIPEA, DCM, 1 h, RT, x 2. *iv*: DIPEA:DCM (10% v/v) 10 min, RT, x 3.

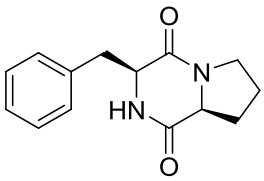
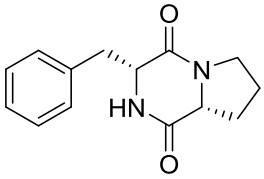
Name	Structure	Scale / mmol	Method (Yield %) Giralt[45]	Method (Yield %) Optimised
cyclo(L-Phe-L-Pro)		0.10	40	-
cyclo(D-Phe-D-Pro)		0.08	-	42

Table 4.2: The yields of a DKP synthesised via the method and conditions quoted in the literature and a DKP synthesised using our optimised strategy.

The syntheses were carried out on a scale that ranged from 0.05 mmol to 0.15 mmol and yields of the isolated and purified DKPs ranged from moderate to excellent (31-86%) (**47**, **54** – **63** in **Figure 4.13**). Overall, the solid syntheses resulted in tens of milligrams of pure product DKP being produced in each case. Additionally, no evidence for the formation of diastereoisomers during the syntheses was observed in any crude product by ^{1}H or ^{13}C NMR.

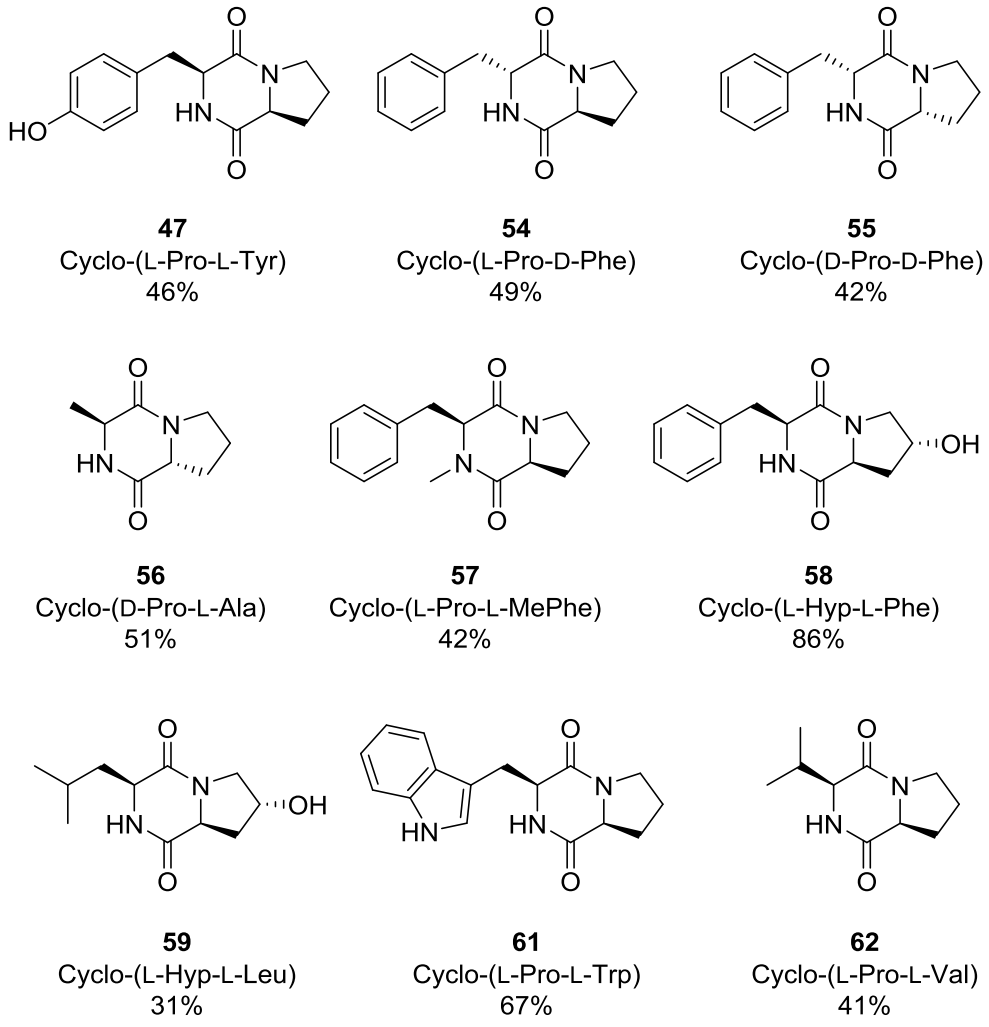


Figure 4.13: The library of DKPs synthesised to be tested against CCL2 induced chemotaxis.

The DKPs were screened for CCL2 induced chemotaxis inhibition (**Section 4.5.2**) and anti-fungal activity (**Section 4.5.1**) alongside commercially available DKPs: Cyclo(Gly-L-Phe) (**60**) and cyclo(Gly-L-Pro) (**64**). Cyclo(L-Phe-L-Pro) (**48**) (synthesised by Dr N. Colgin, **Section 4.2.1**) was also included. This completed the library of target DKPs in **Figure 4.12**.

4.5. Biological Studies of the New Library DKP

4.5.1 CCL2 Induced Chemotaxis: In Vitro

** The following experiments were carried out under the supervision of Dr. C. Barker within the Ali group at Newcastle University Medical School.*

Studies of *in vitro* chemotaxis predominately used a transwell membrane system, in one of two forms: bare membrane chemotaxis (BMC) assay and transendothelial chemotaxis (TEC). In our case, the assays measure induced THP-1 (Human acute monocytic leukemia cell line) cell migration.

The simple BMC (**Figure 4.14**) assay uses a transwell system where medium supplemented with chemokine (± DKP inhibitor) is placed in the lower well and medium supplemented with THP-1 cells is placed in the top well. The chemokine-chemokine receptor interactions cause a cellular response and facilitate the movement of cells from the upper well across a *fenestrated* bare transwell membrane (3 μM pore size) to the lower well where they are counted.

The TEC assay (**Figure 4.15)** follows on from the BMC assay above, however this system involves more biological features: Additions of a monolayer of HMEC-1 (endothelial) cells to the upper aspect of the membrane as well as cytokines IFN-γ and TNF-α to stimulate (up-regulate) adhesion molecules on the HMEC-1 (circles) and THP-1 cells (squares) in **Figure 4.15**.

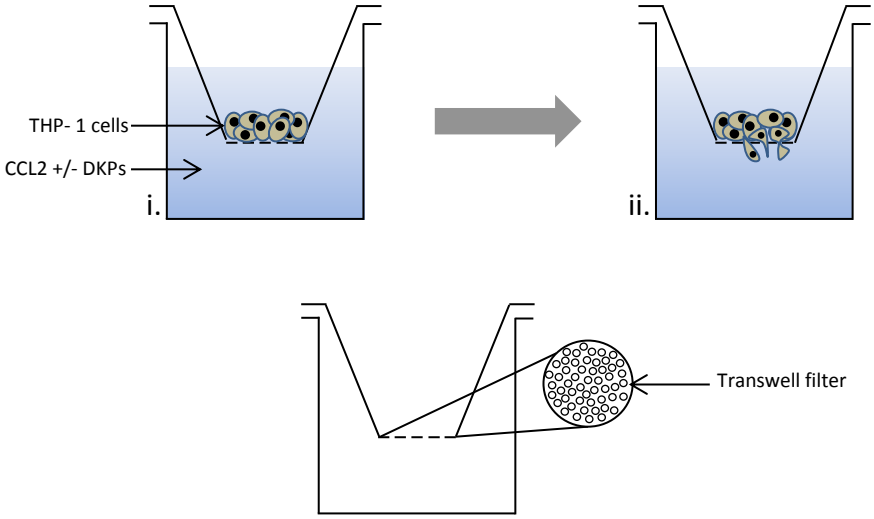


Figure 4.14: A BMC assay. *i*: The lower well serum-free medium supplemented with CCL2 ± DKP. Monocytes in a solution of 1% bovine serum albumin (BSA) medium are added to the upper well. *ii*: The assay was incubated before removal of medium, methanol fixation and haematoxylin staining. The membrane filters dehydrated, mounted and migrant cells counted. [46]

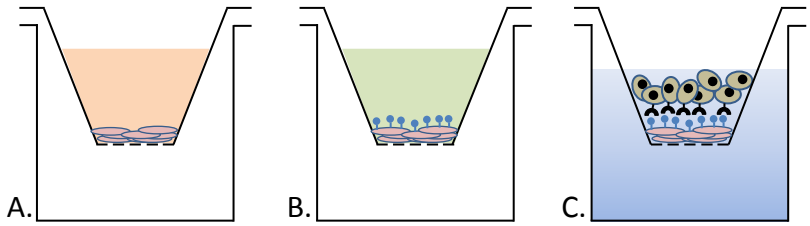


Figure 4.15: A TEC assay A) shows the monolayer of HMEC-1 endothelial cells. B) The stimulation of HMEC-1 cells (circular appendages). C) Introduction of stimulated (square appendages) THP-1 monocytes as well as a chemokine concentration gradient (shading) to promote migration. The staining of cells with haematoxylin enabled counting by high resolution microscopy. The number of cells that had migrated through both the membrane and the endothelial cells was found by counting the cells adherent to the inferior aspect of the membrane.[46]

The TEC assay has increased biological depth and is more applicable to how the compounds would behave in an *in vivo* setting. However, compared to the BMC assay the experiment is more labour intensive, difficult and less economical. Thus, the BMC was predominately used especially for our work i.e. to screen compounds **47**, **48** and **54 - 63**.

4.5.2 Chemotaxis Library Screen via a BMC Assay

In this case we are measuring the ability of the chemokine (CCL2) to interact with chemokine receptors on THP-1 white blood cells (monocytes) and induce chemotaxis. THP-1 monocytes are a cell line derived from the peripheral blood of a 1 year old child with acute monocytic leukemia (ATCC TIB-202). Cells were cultured in complete RPMI-1640. The assay was undertaken using a transwell filter (3 μM) system with 24-well companion plates. Prior to the experiment each companion plate is blocked with 1ml of 1% BSA in RPMI-1640 to prevent unwanted chemokine binding and lowering of effective chemokine concentration in solution.[47,48] The assays were then carried out in triplicate as described in **Figure 4.16** below.

Figure 4.16:A bare membrane chemotaxis assay. *i*: The lower well contains 800 μL of serum-free RPMI-1640 medium supplemented with CCL2 (10 nM) and ± DKP. 500000 monocytes in a 500 μL solution of 1% bovine serum albumin (BSA) in RPMI 1640 were added to the upper well. *ii*: The assay was incubated at 37 °C for 90 min before removal of medium. Non-migrated cells were gently removed from the upper surface of the transwell filter before fixation in cold 100% methanol (1 h) and staining *via* haematoxylin. The filters were dehydrated and mounted to slides. In each filter the migrant cells in five high-power fields of vision were counted.

It was unknown at what concentration of DKP to carry out the library screen as there is significant variation within the structures contained in this library and a lower resemblance to the natural product inhibitor originally tested (**Section 1.4.3**, **1.4.4** and **4.1.2**).[29] Therefore, initially a smaller screen was undertaken using only selected DKPs, **58**, **61** and the previously tested **48** at 100 μM (**Figure 4.17**).

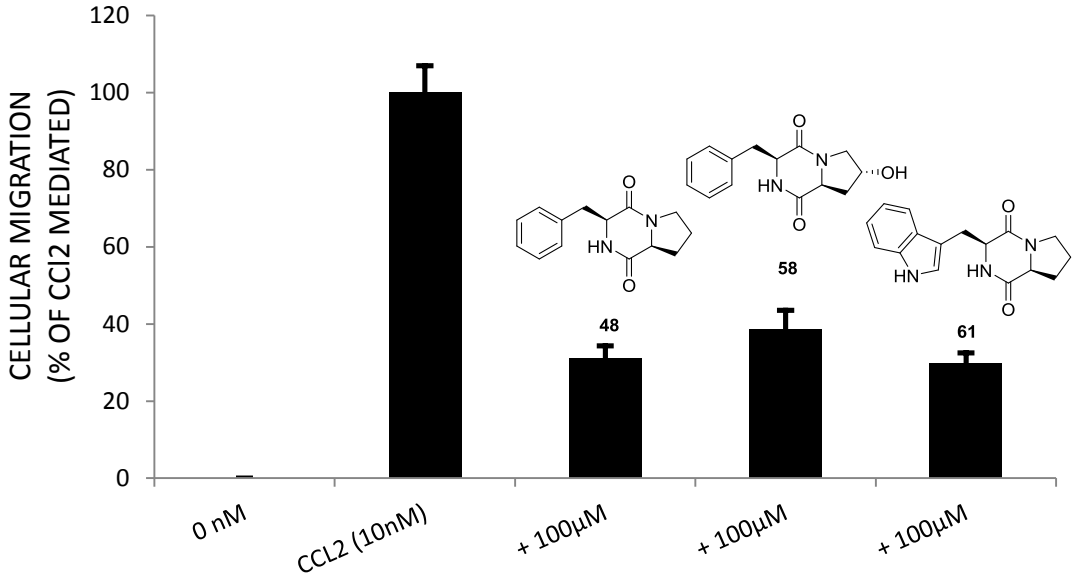


Figure 4.17: Bare membrane chemotaxis assay measuring cellular migration of THP-1 cells in the prescence 10 nM CCL2 and **48**, **58** or **61** at concentration of 100 μM. The control contains no chemokine or DKP. The data is normalised to 100% with 10 nM CCL2 in the absence of inhibitor.

This first screen showed that **58** and **61** were inhibitors of CCl2 induced chemotaxis *in vitro* to a similar extent to **48** and that a 100 μM concentration of DKP was appropriate. Having carried out the initial screen the chemotaxis assay was then repeated on a larger scale with the remainder of the library of DKPs including **48** again as a positive control (**Figure 4.18**).

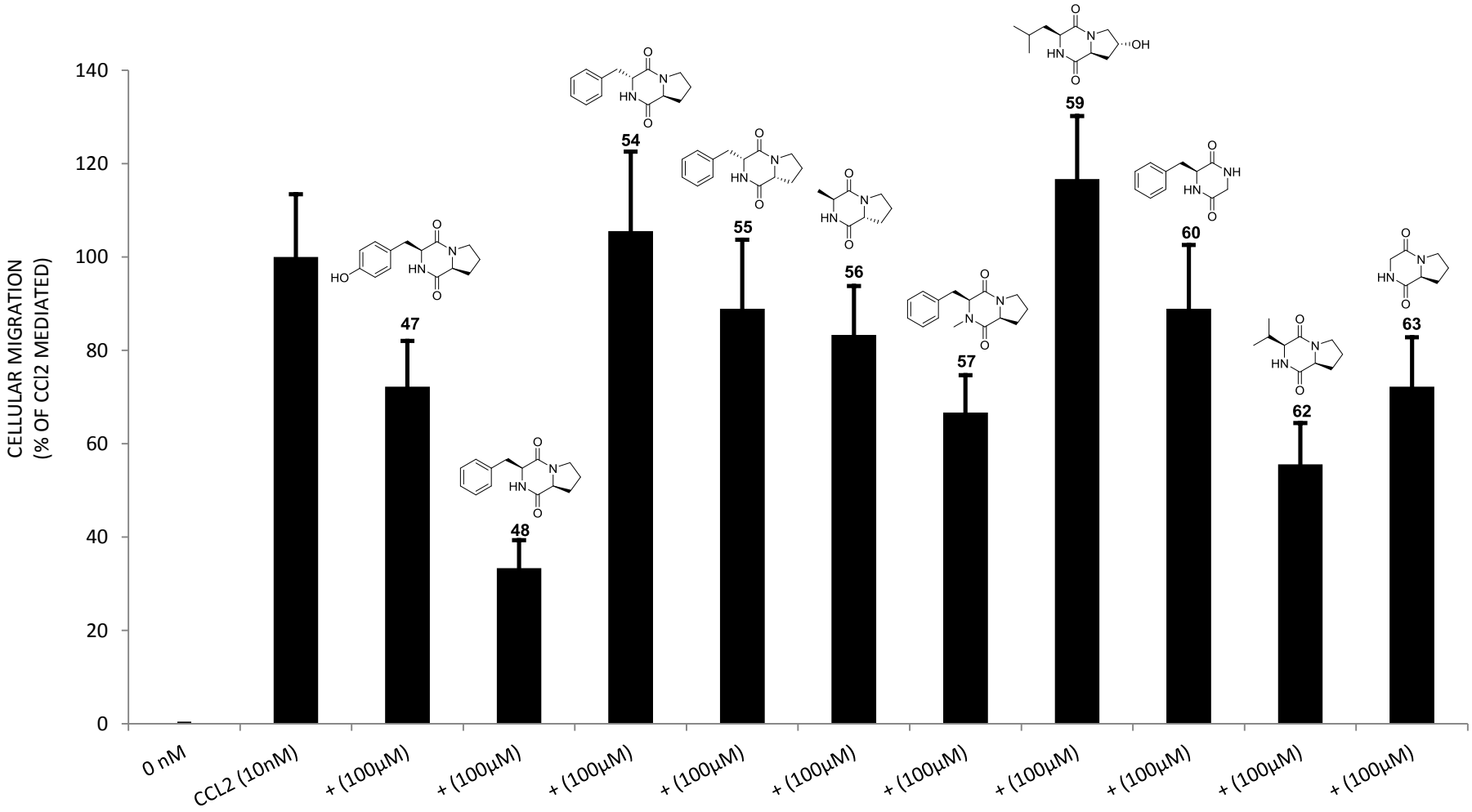


Figure 4.18: Bare membrane chemotaxis assay measuring cellular migration of THP-1 cells in the prescence 10 nM CCL2 and **47**, **48**, **54, 55**, **56**, **57**, **59**, **60**, **62** or **63** at concentration of 100 μM. The control (0 nM) contains no chemokine or DKP. The data is normalised to 100% with 10 nM CCL2 in the absence of inhibitor.

The biological evaluation showed that the majority of compounds (**47**, **54**, **55**, **56**, **57**, **59**, **60**, **62** and **63**) tested are not strong inhibitors of CCL2 induced chemotaxis with cell migration inhibited by only <45% at 100 μM. However, the DKPs **48**, **58** and **61** showed good activity (cell migration inhibited by >62%). The results are of benefit in defining the structural motifs that are of importance. The small library screen showed that both aromatic and prolyl moieties are necessary for significant activity as well as showing the importance of stereochemistry. Preliminary attempts were made to gain X-ray structures of the DKPs synthesised to enable rational explanations to the observed biological results. From the crystallisation experiments only **54** formed crystals of good enough quality to enable structural determinations to be carried out (**Section 4.6**).

4.5.2 Anti-fungal Activity

Bioassays of compounds **47, 48** and **54 – 63** against *C. albicans*, *P. expansum*, *F. culmorum* and *A. flavus* were carried out at University College Dublin (by Dr. N. O'Connor). Although some zones of clearing occurred after 24 h the fungus overgrows the zone of inhibition after 48 h which indicated that the DKPs showed no significant anti-fungal activity.

4.6 Conclusions And Analysis of the DKP Inhibitor Library

4.6.1 Stereochemistry

Previous work did not explore the effect of chirality on inhibition as entirely L-amino acids were tested with the exception of cyclo(L-Pro-D-Tyr) (**50**). **50** showed similar poor inhibition results to cyclo(L-Pro-L-Tyr) (**47**) and thus, no conclusions could be made. In the new library 3 diastereoisomers of the type: cyclo(Pro-Phe) (**48**, **54** and **55**), were screened and noticeably different and relevant results were found. Significant inhibition was not found in the case of (*S*,*R*) cyclo(D-Phe-L-Pro) **54** and cyclo(D-Phe-D-Pro) (*R*,*R*) **55** whereas confirmation of the inhibitory activity of *S*,*S* diastereoisomers **48** was shown. This indicates that the L-L-, motif is important. The crystal structure of **48** and **54** are shown (**Figure 4.19**). **54** adopts a folded structure

where the aromatic ring overlaps the DKP core in a similar way to **47** (**Figure 4.19**). **47** is also known to be a poor chemotaxis inhibitor.

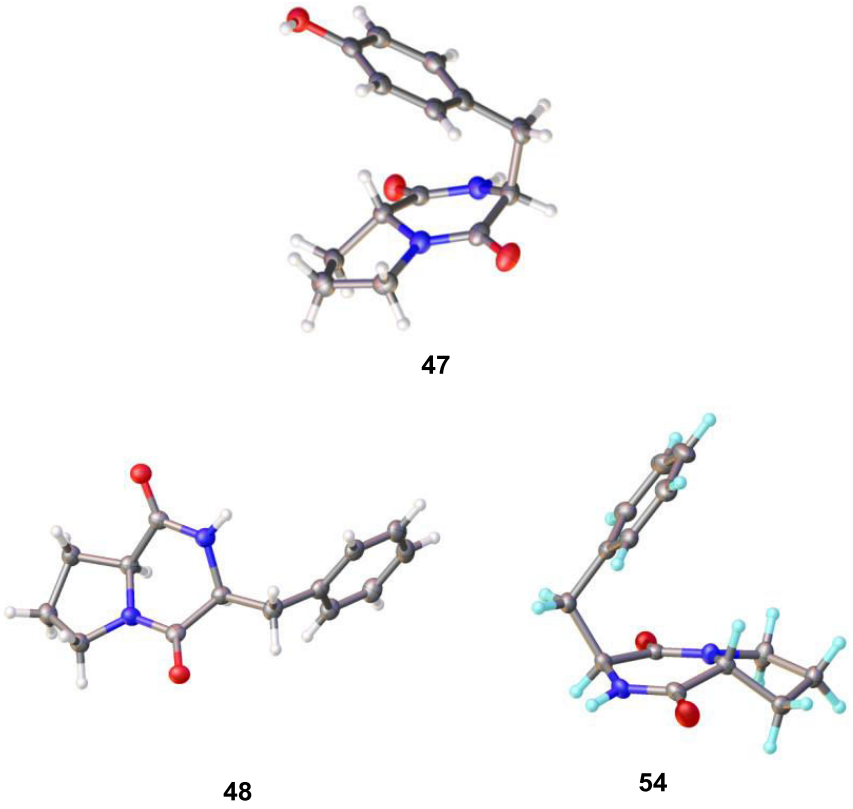


Figure 4.19: Molecular structures of **47***, **48*** and **54** showing 50% probability anisotropic displacement ellipsoids**.** *Crystallised by Mr S. Lear.

4.6.2 Aromatic moiety

The aromatic moiety was found to be very important for biological function as most of the compounds lacking the aromatic moiety (**56**, **59** and **63**) showed minimal inhibition of chemotaxis (cell migration was >75%). Cyclo(L-Val-L-Pro) (**62**) was found to be the most active non-aromatic DKP and was observed to inhibit cellular migration by 45%.

The significant inhibition found in Cyclo(L-Trp-L-Pro) (**61**) is encouraging as it shows tolerance to a larger bicyclic heteroaromatic residue that can have significantly different biological effects to phenylalanine.[49] In our study **61** has comparable activity to **48** at 100 μM. Of note, NMR experiments conducted by Grant

et al.[50] showed **61** to have preference for an elongated structure in $CDCl_3$ (71%) and a folded structure in d_6-DMSO (52%). X-ray crystallography observed the lowest energy conformations to be in a folded state.[50] Thus, future libraries must synthesise and test a broader scope that includes multi-cyclic aromatic and heteroaromatic moieties. Subsequent tests at lower concentrations would need to be carried out to determine if **48** or **61** is the more potent inhibitor of CCL2 induced chemotaxis.

4.6.3 Proline moiety

The proline moiety is vital to function as no inhibition is observed when this group is either removed cyclo(Gly-L-Phe) (**60**) or the opposite stereochemistry is used (cyclo (D-Pro-L-Ala) (**56**) and cyclo(D-Pro-D-Phe) (**55**). Cyclo(L-Phe-L-Hyp) (**58**) showed significant inhibition to almost the same extent as **61** and **48**. Even though the activity is slightly diminished this result shows the tolerance to a functional group in the 4-position of the 5-membered prolyl ring and perhaps substituents other than a hydroxyl could have a positive effect.

4.7 Library Expansion Using Unnatural Aromatic and Heteroaromatic Amino Acids

4.7.1 Target Molecules

Chapter 2 and **3** describe the syntheses of numerous unnatural aromatic and proline derivative amino acids with protecting group strategies that can easily translate to the synthesis of DKPs. Thus, the synthesis of a diverse selection of DKPs could be envisaged (**Figure 4.20**).

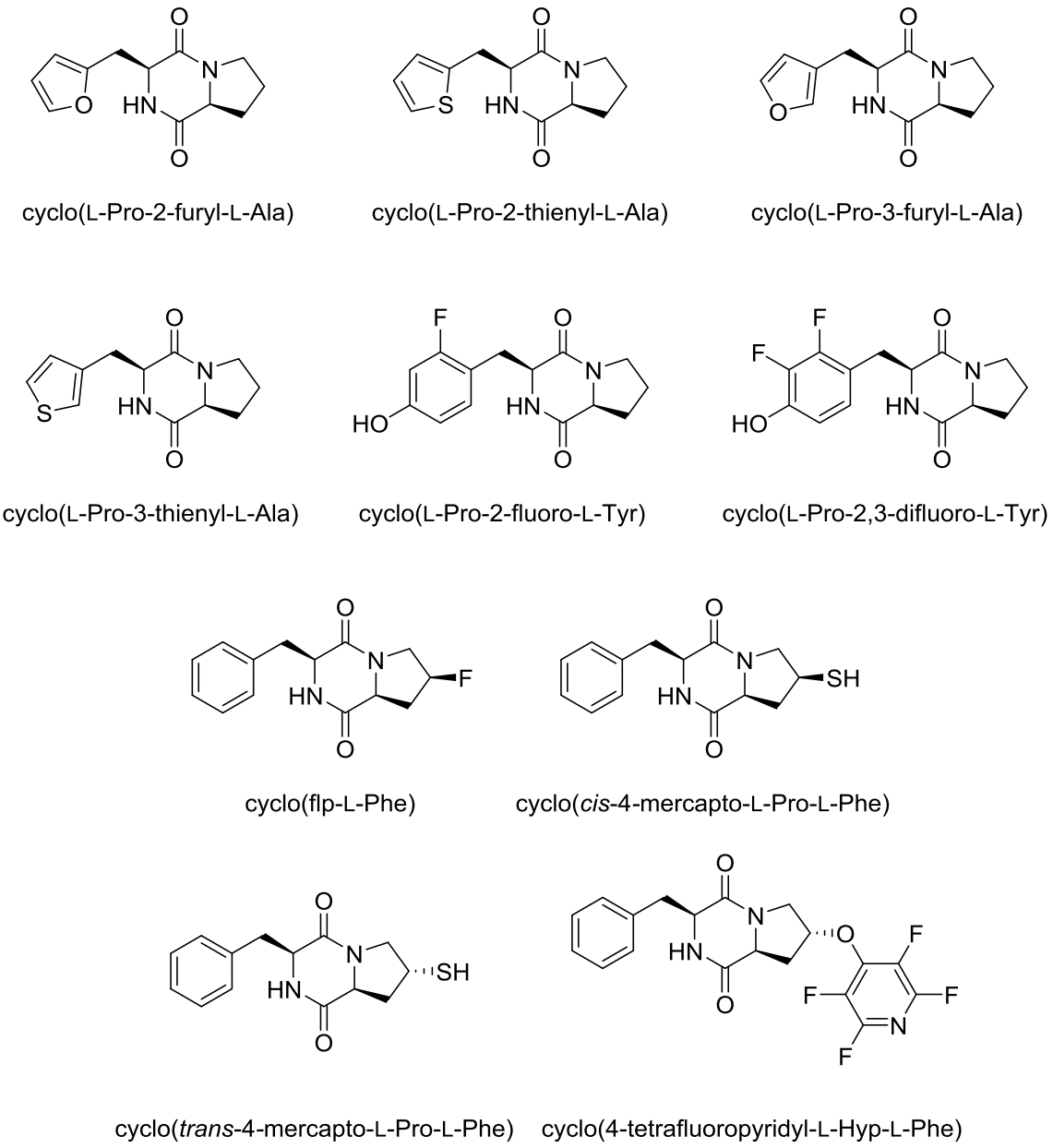
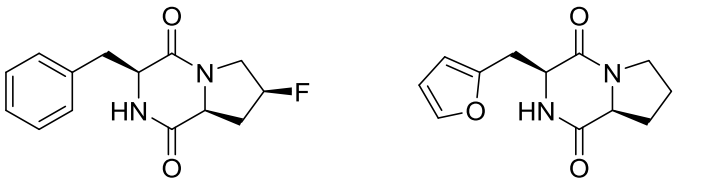


Figure 4.20: Selection of DKP molecules that could be synthesised from unnatural amino acids prepared in **Chapter 2** and **3**.

The testing of compounds of this type is beyond the scope of this thesis. However, a route to their synthesis would be beneficial in order to show the trivial nature of unnatural amino acid incorporation into a DKP. An aromatic derivative (2-furylalanine) from **Chapter 2** and a prolyl derivative (4-fluoroproline) from **Appendix 1** were utilised.

4.7.2 Syntheses of Cyclo(flp-L-Phe) and Cyclo(L-Pro-2-furyl-L-Ala)

The two novel target molecules are cyclo(flp-L-Phe) (**66**) and cyclo(L-Pro-2-furyl-L-Ala) (**67**) (**Figure 4.21**)

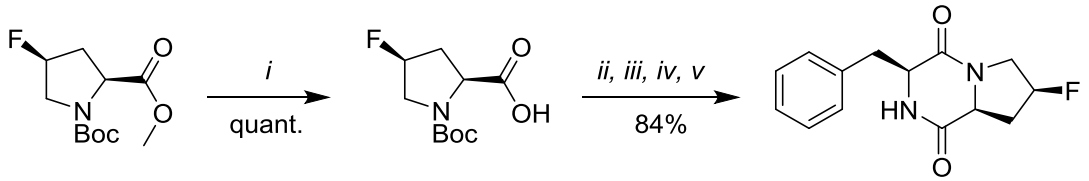


66
cyclo(flp-L-Phe)

67
cyclo(L-Pro-2-furyl-L-Ala)

Figure 4.21: Structures of novel target DKP molecules **66** and **67**.

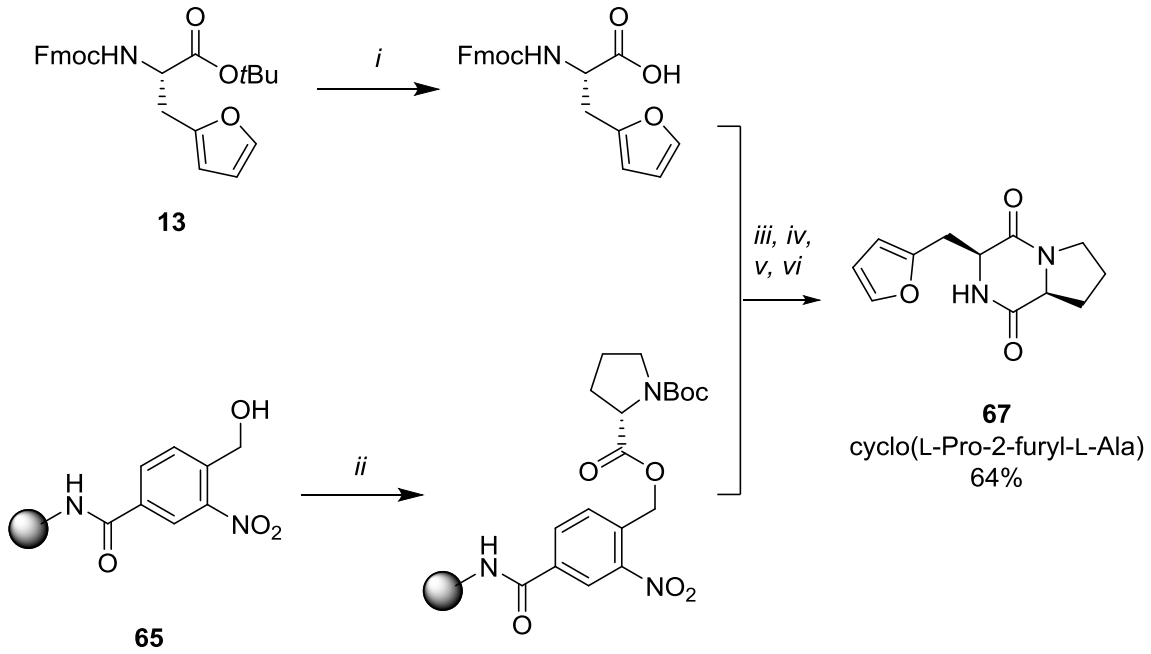
Our modified Giralt[45] strategy of on-resin synthesis of DKPs utilised *N*-Boc amino acids. The synthesis of Boc-2-furyl-L-Ala-OBn (**19**) and Boc-*cis*-fluoroproline-OMe (**68**) is described in **Chapter 2** and **Appendix 1**, respectively. Deprotection of the carboxy terminal protection would yield compounds ready for the on-resin synthesis. The basic deprotection of **68** yields the Boc-flp-OH (**69**) in quantitative yield and subsequent DKP formation was carried out in 84% yield (**Scheme 4.8**) over 5 steps.



68 **69** **66**
cyclo(flp-L-Phe)

Scheme 4.8:*i*: LiOH, H_2O, THF, RT, 18 h. : *ii*: DIC, DMAP, DCM, 30 min, RT, x 2. *iii*: TFA:DCM (40% v/v), 10 min, RT, x 2. *iv*: Boc-Phe-OH, PyBOP, DIPEA, DCM, 1 h, RT, x 2. *iv*: TFA:DCM (40% v/v), 10 min, RT, x 2 *v*: DIPEA:DCM (10% v/v) 10 min, RT, x 3.

The benzyl deprotection of **19** is problematic (**Chapter 2**, **Section 2.3.6)** (**Scheme 4.9**) and the Giralt group showed precedence for utilising *N*-Fmoc amino acids in their DKP synthesis strategy.[45] Therefore, we envisaged utilising Fmoc-2-furyl-L-Ala-O*t*Bu (**13**) with acidic O*t*Bu deprotection. **13** was deprotected (to yield **70)** and successfully utilised in a slightly modified on-resin synthesis yielding product DKP: **67** in 64% yield (relative to the resin) over 4 steps (**Scheme 4.9**)



Scheme 4.9: Synthesis of **67**. *i:* TFA:DCM (40 % v/v), 30 min *ii*: Boc-Pro-OH, DIC, DMAP, DCM, 30 min, RT, x 2. *ii*: TFA:DCM (40% v/v), 10 min, RT, x 3. *iii*: PyBOP, DIPEA, DCM, 1 h, RT, x 2.i*v*: piperidine:DMF (20% v/v) 10 min, RT, x 2.

4.8 General Conclusions

The on-resin strategy used to synthesize a library of DKPs was successful as it enabled the rapid production of pure compound in moderate to excellent yield from simple Boc-amino acids. The strategy provided significant benefits over the original solution phase approach (**Section 4.2.1**) e.g. less purification steps, lowering reaction temperature and lowering total time.

The only noticeable disadvantage is that solid-supported syntheses are generally less transferable to scale up. However, in our case the multi-milligrams of pure product were sufficient to carry out all screening.

The biological testing provided two inhibitors (**58** and **61**) to match inhibitors from the previous study (**46**, **48** and **49**). However, due to its broader scope of DKPs it has helped the narrow substrate scope needed to retain activity with the variety of inactive compounds produced. Moreover, the activity of **61** and **58** has opened avenues to investigate with regards to heteroaromatic L-amino acids and substituted L-prolines, respectively. Work to determine accurate IC_{50} values is ongoing.

4.9 References

1. M. B. Martins and I. Carvalho, *Tetrahedron*, 2007, **63**, 9923-9932.
2. A. D. Borthwick, *Chem. Rev.*, 2012, **112**, 3641-3716.
3. P. Belin, M. Moutiez, S. Lautru, J. Seguin, J.-L. Pernodet and M. Gondry, *Nat. Prod. Rep.*, 2012, **29**, 961-979.
4. H. Kanzaki, D. Imura, T. Nitoda and K. Kawazu, *J. Biosci. and Bioeng.*, 2000, **90**, 86-89.
5. S. Lautru, M. Gondry, R. Genet and J.-L. Pernodet, *Chemi. Biol.*, 2002, **9**, 1355-1364.
6. C. D. Ericsson, H. L. DuPont, P. Sullivan, E. Galindo, D. G. Evans and D. J. Evans, Jr., *Ann. Intern. Med.*, 1983, **98**, 20-25.
7. K. McCleland, P. J. Milne, F. R. Lucieto, C. Frost, S. C. Brauns, M. Van De Venter, J. Du Plessis and K. Dyason, *J. Pharm. Pharmacol.*, 2004, **56**, 1143-1153.
8. D. R. Houston, B. Synstad, V. G. H. Eijsink, M. J. R. Stark, I. M. Eggleston and D. M. F. van Aalten, *J. Med. Chem.*, 2004, **47**, 5713-5720.
9. P. L. Rodriguez and L. Carrasco, *J. Virol.*, 1992, **66**, 1971-1976.
10. G. Kilian, H. Jamie, S. C. Brauns, K. Dyason and P. J. Milne, *Pharmazie*, 2005, **60**, 305-309.
11. P. Klausmeyer, O. M. Z. Howard, S. M. Shipley and T. G. McCoud, *J. Nat. Prod.*, 2009, **72**, 1369-1372.
12. O. S. Kwon, S. H. Park, B. S. Yun, Y. R. Pyun and C. J. Kim, *J. Antibiot.*, 2000, **53**, 954-958.
13. A. Mendham, J. Spencer, B. Chowdhry, T. Dines, M. Mujahid, R. Palmer, G. Tizzard and S. Coles, *J. Chem. Crystallogr.*, 2011, **41**, 1323-1327.
14. P. M. Fischer, *J. Pept. Sci.*, 2003, **9**, 9-35.
15. R. Palmer, B. Potter, A. Mendham, T. Dines and B. Chowdhry, *J. Chem. Crystallogr.*, 2010, **40**, 608-615.
16. R. B. Corey, *J. Am. Chem. Soc.*, 1938, **60**, 1598-1604.
17. J. D. Hirst and B. J. Persson, *J. Phys. Chem. A*, 1998, **102**, 7519-7524.
18. M. Nakao, Y. Toriuchi, S. Fukayama and S. Sano, *Chem. Lett.*, 2014, **43**, 340-342.
19. Z. Y. Li and S. Mukamel, *J. Phys. Chem. A*, 2007, **111**, 11579-11583.
20. A. D. Borthwick and J. Liddle, *Retosiban and Epelsiban: Potent and Selective Orally Available Oxytocin Antagonists*, Wiley-VCH Verlag GmbH & Co. KGaA, 2013.
21. D. D. Long, R. J. Tennant-Eyles, J. C. Estevez, M. R. Wormald, R. A. Dwek, M. D. Smith and G. W. J. Fleet, *J. Chem. Soc., Perkin Trans. 1*, 2001, 807-813.
22. A. C. Mita, R. S. Heist, O. Aren, P. N. Mainwaring, L. Bazhenova, S. M. Gadgeel, R. H. Blum, J. Polikoff, J. Biswas and M. A. Spear, *J. Clin. Oncol.*, 2010, **28**, 1.
23. M. Millward, P. Mainwaring, A. Mita, K. Federico, G. K. Lloyd, N. Reddinger, S. Nawrocki, M. Mita and M. A. Spear, *Invest. New Drugs*, 2012, **30**, 1065-1073.
24. B. Nicholson, G. K. Lloyd, B. R. Miller, M. A. Palladino, Y. Kiso, Y. Hayashi and S. T. Neuteboom, *Anti-cancer drugs*, 2006, **17**, 25-31.

25. K. Kanoh, S. Kohno, T. Asari, T. Harada, J. Katada, M. Muramatsu, H. Kawashima, H. Sekiya and I. Uno, *Bioorg. Med. Chem. Lett.*, 1997, **7**, 2847-2852.
26. M. D. Turner, B. Nedjai, T. Hurst and D. J. Pennington, *Biochim. Biophys. Acta-Mol. Cell Res.*, 2014, **1843**, 2563-2582.
27. R. Horuk, *Nat. Rev. Drug Discov.*, 2009, **8**, 23-33.
28. D. J. Fox, J. Reckless, H. Lingard, S. Warren and D. J. Grainger, *J. Med. Chem.*, 2009, **52**, 3591-3595.
29. M. Saleki, N. Colgin, J. A. Kirby, S. L. Cobb and S. Ali, *MedChemComm*, 2013, **4**, 860-864.
30. M. Jainta, M. Nieger and S. Bräse, *Eur. J. Org. Chem.*, 2008, **2008**, 5418-5424.
31. V. A. Basiuk, T. Y. Gromovoy, A. A. Chuiko, V. A. Soloshonok and V. P. Kukhar, *Synthesis*, 1992, **1992**, 449-451.
32. K. C. Majumdar, K. Ray and S. Ganai, *Synlett*, 2010, 2122-2124.
33. E. Fischer, *Ber. Bunsen-Ges.*, 1906, **39**, 2893-2931.
34. S. Lee, T. Kanmera, H. Aoyagi and N. Izumiya, *Int. J. Pept. Protein Res.*, 1979, **13**, 207-217.
35. K. Suzuki, Y. Sasaki, N. Endo and Y. Mihara, *Chem. Pharm. Bull.*, 1981, **29**, 233-237.
36. L. Perez-Picaso, J. Escalante, H. F. Olivo and M. Y. Rios, *Molecules*, 2009, **14**, 2836-2849.
37. G. Cabon, B. Gaucher, A. Gegout, S. Heulle and T. Masquelin, *Chimia*, 2003, **57**, 248-254.
38. C. Hulme and V. Gore, *Curr. Med. Chem.*, 2003, **10**, 51-80.
39. Q. Lin and H. E. Blackwell, *Chem. Commun.*, 2006, 2884-2886.
40. C. Hulme, M. M. Morrissette, F. A. Volz and C. J. Burns, *Tetrahedron Lett.*, 1998, **39**, 1113-1116.
41. H. Habashita, M. Kokubo, S. Hamano, N. Hamanaka, M. Toda, S. Shibayama, H. Tada, K. Sagawa, D. Fukushima, K. Maeda and H. Mitsuya, *J. Med. Chem.*, 2006, **49**, 4140-4152.
42. J. J. Chen, A. Golebiowski, S. R. Klopfenstein and L. West, *Tetrahedron Lett.*, 2002, **43**, 4083-4085.
43. J. P. Bourgault, A. R. Maddirala and P. R. Andreana, *Org. Biomol. Chem.*, 2014, **12**, 8125-8127.
44. J. C. O'Neill and H. E. Blackwell, *Comb. Chem. High T. Scr.*, 2007, **10**, 857-876.
45. M. Teixido, E. Zurita, M. Malakoutikhah, T. Tarrago and E. Giralt, *J. Am. Chem. Soc.*, 2007, **129**, 11802-11813.
46. M. Baggiolini and P. Loetscher, *Immunol. Today*, 2000, **21**, 418-420.
47. M. De Buck, M. Gouwy, P. Proost, S. Struyf and J. Van Damme, *Biochem. Pharmacol.*, 2013, **85**, 789-797.
48. S. Struyf, P. Proost, J. P. Lenaerts, G. Stoops, A. Wuyts and J. Van Damme, *Blood*, 2001, **97**, 2197-2204.
49. P. Braun and G. von Heijne, *Biochemistry*, 1999, **38**, 9778-9782.
50. G. D. Grant, A. L. Hunt, P. J. Milne, H. M. Roos and J. A. Joubert, *J. Chem. Crystallogr.*, 1999, **29**, 435-447.

Chapter 5 : Attempted Syntheses of the Native Human Chemokine CCL2 and its Analogues

5.1 Protein Synthesis

5.1.1 Chemical Synthesis of Peptides/Proteins

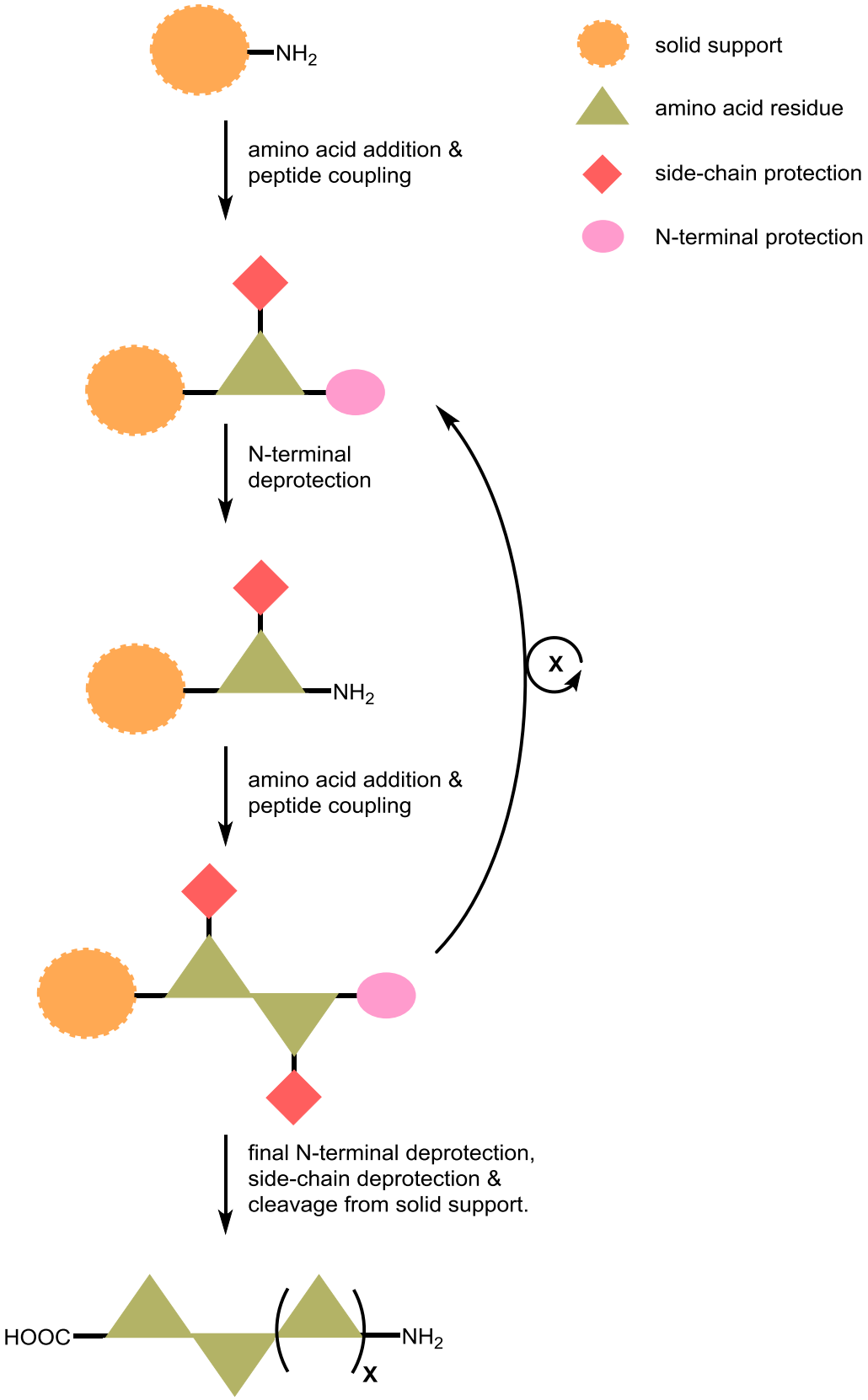
The total chemical synthesis of a protein is the complete synthesis of a protein from the constituent amino acids. Originally, strategies utilised fully protected peptides in solution phase syntheses and these early approaches resulted in the production of a number of bioactive molecules.[1] For example, an active oxytocin-amide octapeptide amide by du Vigneaud in 1953.[1] However, severe solubility, purification and characterisation limitations (of the fully protected segments) hampered development and narrowed the scope of accessible peptide target molecules.[2]

A significant advancement came with the advent of solid phase peptide synthesis (SPPS) pioneered by Merrifield in the 1960s. SPPS provided a powerful tool to enable the total synthesis of fully active proteins. A landmark example in the field was Merrifield's *linear stepwise* synthesis of ribonuclease A (1971, 124 amino acids), work for which he later won the Nobel prize.[3] The boundaries were further advanced with the use of native chemical ligation (NCL) pioneered by Kent in the 1990s. The total chemical synthesis of proteins of over 200 amino is now possible, as evidenced by the synthesis of a HIV-1 protease enzyme (203 amino acids).[4]

Total chemical synthesis gives chemists the precise control needed to finely tune bioactive peptides and proteins by incorporation of unnatural amino acids, labels and chelates. Chemical synthesis gives the freedom to change steric or electronic factors and to tailor the size, nucleophilicity, acidity, hydrophobicity and/or hydrogen bonding character of a specific residue, in a selective fashion.[5]

5.1.2 SPPS of Peptides/Proteins

Generally, SPPS utilises N-amino protected amino acids with orthogonally protected side chains and a solid support (resin) in a sequence of repetitive reactions (**Scheme 5.1**).[6,7] The sequence of reactions starts with coupling an amino acid to the linker and resin and then subsequent repetitions of N-amino deprotection before coupling the next residue. Finally, cleavage of the solid support and any amino acid side chain protection yields the full peptide.[8]



Scheme 5.1: A simplified SPPS strategy to chemically synthesise a peptide. The series of repetitive (addition, coupling and deprotection) steps that build from C to N-terminus are shown.[6]

The crude peptide product produced *via* SPPS should be relatively pure, but a number of impurities can arise from incomplete peptide couplings, side-chain deprotection and unwanted side-reactions. Typically, RP-HPLC is used to obtain a homogenous mixture of the desired pure peptide after the synthesis.

The length of a peptide that can be synthesised by SPPS is variable and it is highly sequence dependant. Problems with efficient peptide couplings and side chain deprotection increase in peptides longer than 40 amino acids and in many cases there is a significant drop in crude purity beyond 50 amino acids. In order to access peptides of longer that 50 amino acids a number of strategies have been developed to overcome the problems and ultimately push the scope of SPPS. Of these, NCL (**Chapter 3**) is the most widely used technique in the total syntheses of proteins.[9]

5.1.3 Native Chemical Ligation (NCL)

In the NCL reaction an unprotected peptide containing an N-terminal cysteine reacts with a peptide chain containing a C-terminal thioester to yield a natural amide bond at the linkage site (junction).[2] NCL relies on the chemoselective coupling of two unprotected peptides in a neutral aqueous solution, one bearing a C-terminal thioester and the other an N-terminal cysteine. Therefore, to reach a synthetic target there must be a suitable placed cysteine to form a ligation site as well as the non-trivial production of a C-terminal thioester containing peptide.[10]

5.1.4 Thioester and Solid Support

In terms of SPPS, the mechanistic and catalytic intricacies of the thioester and thiol leaving group are important. Predominantly alkyl thioesters are used and they are prepared *via* direct synthesis using different types of thioester-linked resin. For example, the sulfamylbutyryl was the first developed thioester resin (for Boc-SPPS[11] and later for Fmoc-SPPS) and it has facilitated the total synthesis of a variety of long peptides/proteins. The thioesters are stable to cleavage conditions, handling/purification as well as being fairly unreactive under NCL conditions. This low reactivity of the alkyl thioesters initially formed after SPPS means an exogenous thiol catalyst is normally

added to create an *in situ* "active thioester" species with superior reactivity. E.g. the aryl thiol 4-mercaptophenylacetic acid (MPAA).[12] MPAA has become the thiol catalyst of choice due to it activity, water solubility and low odour.[12] However, alky thioester resins have a number of disadvantages, including: low yields, lack of facile (test-cleavage) reaction monitoring and the use of dangerous alkylating reagents (diazomethane or iodoacetonitrile).[13,14]

Very recently thioester "mimic" resins (Dawson Dbz and bis(2-sulfanylethyl)amino (SEA)) have been developed to overcome the aforementioned problems and offer different properties to the sulfamylbutyryl resins. However, literature examples of both thioester and thioester "mimic" resins being utilised in microwave assisted SPPS are still relatively few. A list of Fmoc SPPS-compatible "activated" resins and their descriptions are given in **Table 5.1**.

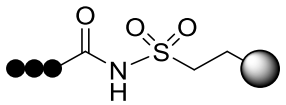
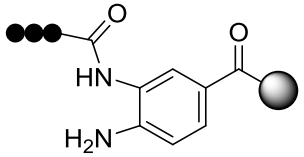
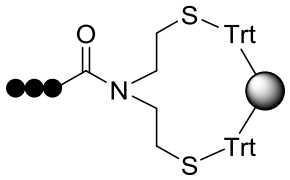
Name	Structure	Commercial Availability (Novabiochem® / Iris Biotech ©)	Notes	Microwave Assisted SPPS ref(s)	
				Literature	Commercial
Sulfamylbutyryl		-Loading difficulties. Thus, mainly available in pre-loaded form. -varied polymer swelling and loading properties: high-loading polystyrene (PS) and low-loading hybrid PEG/PS). - Boc- and Fmoc-SPPS compatible.	On resin alkylation and nucleophilic thiol cleavage gives rise to low yields and prevents facile synthesis monitoring *via* test-cleavage. However, recently new pre-loaded (Ala and Gly) hybrid-linker resins (sulfamylbutyryl – rink amide) are available. The hybrid link enables facile test-cleavage and removes the need for nucleophilic thiol cleavage.	Elsawy *et al.*[15]	CEM corp.[16]
Dawson Dbz		-Facile, yet, residue specific loading. Therefore, resin available unloaded. -varied polymer swelling/loading properties: high-loading polystyrene (PS) and low-loading hybrid PEG/PS)	Chain branching occurs in some cases, especially with glycine rich sequences. In these cases, pre-synthesis Alloc protection of the non-propagating aromatic amine prevents branching.	Gunasekera *et al.*[17]	-
SEA		-Recently released, no resin known loading issues. Resin available unloaded and pre-loaded. -Only available as low-loading PS	Post-TFA cleavage the peptide C-terminus undergoes reversible N-, S-acyl shift. However, reversibility is prevented (by oxidation of the dithiol group) to simplify purification and provides long-term stability.	-	-

Table 5.1: Properties of sulfmylbutryryl resin as well as thioester "mimic" resins (Dawson Dbz and SEA). Their use in microwave synthesis is referenced.

5.2 Previous Syntheses of Human Chemokine CCL2

CCL2 is a protein of 76 amino acids in length and it contains hydrophobic regions prone to aggregation and secondary structure formation (a "long and difficult" sequence) (**Figure 5.1**). Thus, the development of a route to the total chemical synthesis and purification of CCL2 is non-trivial.

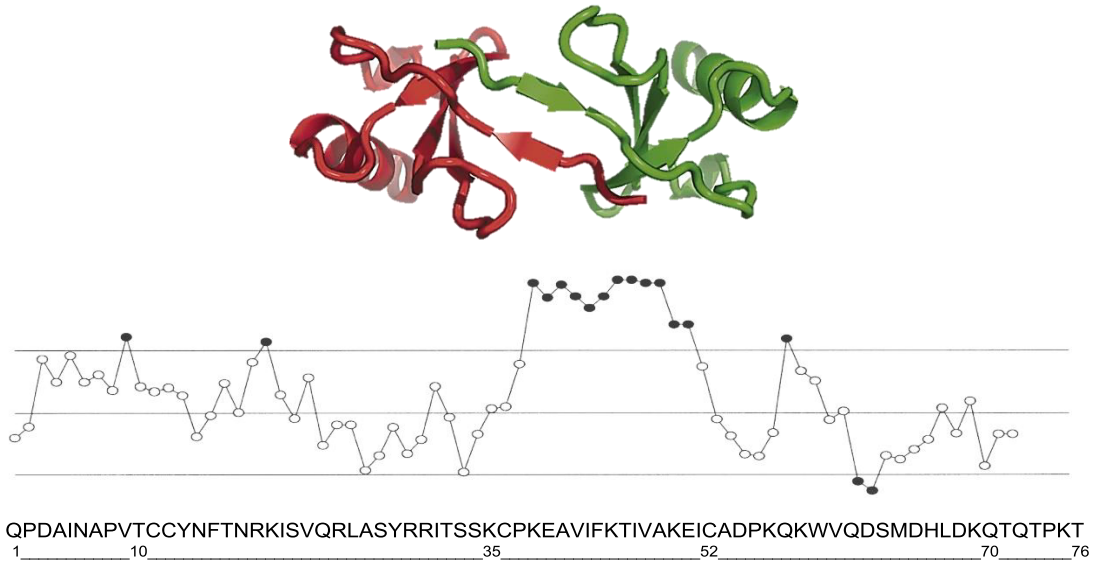


Figure 5.1: *Top*: X-ray crystal structure of CCL2 dimer with asymmetric monomer units indicated (red and green).[9] *Middle*: the "aggregation potential/difficult sequences" graph from Peptide Companion. This is a visual representation of a set of defined mathematical rules[18] which assign additive difficulty values to individual amino acid couplings. This graph highlights the hydrophobic region 39-51 as very difficult. *Bottom*: The primary structure of CCL2 written from N-C terminus.

An automated room temperature synthesis strategy was successfully completed in 1995 by Gong and Clark-Lewis[19] utilising *N*-Boc protection chemistry. One year later an Fmoc based strategy was developed on an automated room temperature synthesiser of the same model by Brown *et al.*[20] Since then the development of other total SPPS routes has allowed access to a host of biologically relevant analogues as well as the native chemokine. CCL2 is therefore a suitable target on which to develop and test new SPPS and purification methods and reagents designed to enable the total chemical synthesis of proteins and protein analogues (e.g. nitrated chemokines). To date, no microwave assisted SPPS of CCL2 have been reported.

5.2.1 The First Total Chemical Synthesis of CCL2

In 1995 Gong and Clark-Lewis published the first synthesis of native CCL2 as well as the synthesis of analogues containing a variety of N-terminal modifications.[19] Chemotaxis assays showed that the N-terminal domain was vital for protein function and specific emphasis was put on the integrity of residues 1-6 to retain functional activity and residues 7-10 for receptor desensitisation but not function.[19] The Boc-based automated linear synthesis had previously been applied to the synthesis of chemokines CXCL7 and CXCL8.[19,21] However, no specific synthetic details of any of these chemokines were given and no difficulties were highlighted. Native CCL2 and the N-terminal analogues folded spontaneously (in the absence of free thiols) and were purified *via* RP-HPLC with a retention time characteristic of other chemokines.[21,22] The crude synthetic yield was stated as 99.3% per amino acid addition, yet the overall purified isolated yield was not given.

As previously mentioned, *N*-Boc SPPS is unpopular and often unfeasible due to the repetitive use of strong acids (TFA) as well as the risks involved in the handling and usage of liquid HF required for resin cleavage. Thus, the majority of CCL2 syntheses have focused on *N*-Fmoc approaches.

5.2.2 A Linear Stepwise N-Fmoc Total Chemical Syntheses of CCL2

In 1996 the first *N*-Fmoc SPPS of CCL2 was published by the Welch group.[20] The automated synthesis was carried out on a Wang resin (resin loading not provided) using an automated ABI room temperature synthesiser fitted with a UV monitor*. The strategy highlighted that acetyl-capping and a double coupling step of every residue (excluding glycine).[20] An acid stable/base labile (tetrabenzo[a,c,g,i] fluorenyl-17 methoxycarbonyl (Tbfmoc), **Figure 5.2**) protecting group/probe was incorporated to the N-terminus pre-cleavage. This hydrophobic protection enables RP-HPLC and graphitized carbon affinity chromatography to give fast, simple and effective purification. The mild removal of the Tbfmoc yields the native (unfolded) chemokine (<1% yield).[20]

*A UV monitor fitted to an automated synthesiser enables the user to track residue by residue Fmoc removal extent and can be used to help to define problematic regions in the synthesis.

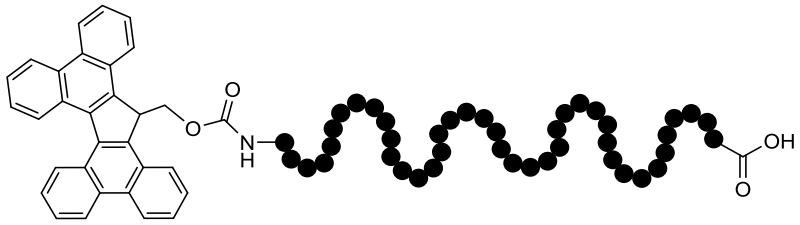


Figure 5.2: N-Tbfmoc protected polypetide.

Two *N*-Fmoc based linear stepwise syntheses of CCL2 were published in 2006 by Kruszynski *et al.*[23,24] and Reid *et al.*[25]

The SPPS published by Reid *et al.* used the product CCL2 in co-crystallisation with the inhibitory antibody: 11k2. This led to the key residues in antibody-CCL2 binding being mapped and pinpointed to antibody-Phe101 insertion into a hydrophobic pocket of CCL2.[25] Although this binding does not extensively overlay the receptor binding site the close proximity leads to significant steric interactions that prevent receptor binding.

The Kruszynski SPPS aimed to produce biotin tagged chemokine that still retains natural function. The synthesis of native CCL2 was carried out using 3 methods on 2 different room temperature automated synthesisers. Surprisingly, it was found that the method that utilised double coupling of amino acids (DCC/HOBt) as standard with additional capping steps on an ABI; model 431A led to considerably lower purity of crude product peptide[23] than the other two methods that were carried using single couplings on either a Symphony Multiple Peptide Synthesizer SMPS-110 or the aforementioned ABI; model 431A. The product CCL2 was purified (by RP-HPLC), oxidised and re-purified by affinity chromatography *via* a monoclonal antibody (mouse anti-human CCL2 monoclonal antibody (mAb)). Effective post-synthesis biotin modification of the synthetic whole protein proved difficult as it led to the indiscriminate biotinylation of residues important for function lysine residues and significantly reduced the derivatives activity.[23] Therefore, follow-up total syntheses site-selectively incorporated an Fmoc-Lys(Alloc)-OH building block at

less "vital to function" Lys69 or Lys75.[†] *N*-Alloc side-chain deprotection and subsequent on-resin biotinylation yielded site-specifically modified CCL2 derivatives that retained activity comparable to the native chemokine (**Figure 5.3**).[24]

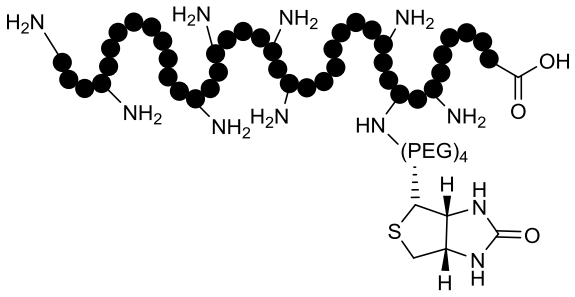


Figure 5.3: Site selectively biotin (with PEG_4 linker) labelled polypeptide.

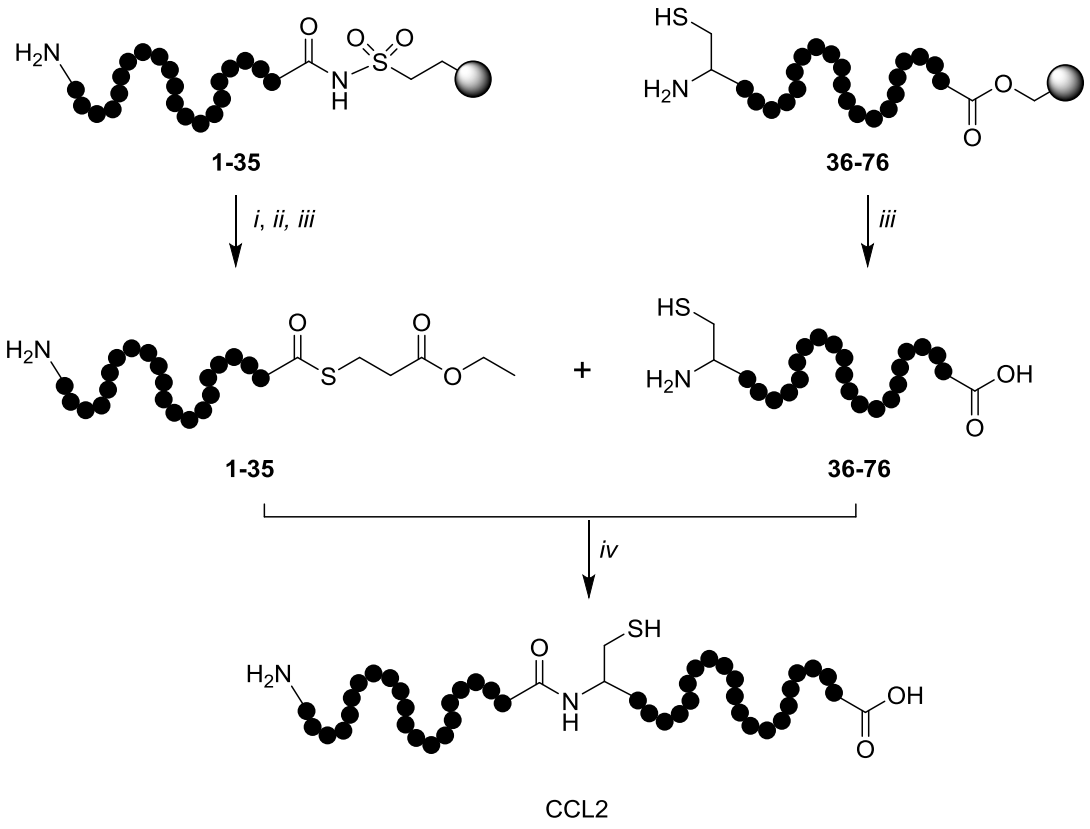
5.2.3 Total Chemical Synthesis of CCL2 *via* NCL

Recently the synthesis of CCL2 *via* a segmented NCL approach of two peptide segments (AAs **1-35** and **36-76**, **Scheme 5.2**) was reported[‡].[9] This strategy also utilised seven evenly spaced pseudo-proline[§] dipeptide building blocks to disrupt the aggregation of growing peptide chains that leads incomplete coupling steps.[26] The C-terminal segment (AA **1–35**) was synthesised at room temperature on a low-loading Wang resin (0.12 mmol /g). The N-terminal (AA **36–76**) segment was synthesised at room temperature on a sulfamylbutryl resin so as to yield a C-terminal thioester. Post-synthesis alkylation of the thioester resin (with (trimethylsilyl)diazomethane ($TMS\text{-}CHN_2$)) followed by thiol attack released the C-terminal thioester from the solid support (**Scheme 5.2**). The peptide segments were purified before the NCL reaction to yield native CCL2.

[†]Native CCL2 contains nine lysine residues and Lys69/Lys75 were chosen in an attempt to minimize disruptions in the bioactivity. Lysine residues: 35, 38 and 49 are known to be important to activity and Lys 19, 44, 56 and 58 are closer to the active surface.

[‡] This NCL approarch was reported in 2010, as work in the Cobb group was being carried out.

[§]Pseudo-prolines are serine, threonine and cysteine derivative building blocks with an acid lable cyclic side-chain protection. The side-chain protection is in the form of a 5-membered heteroaromatic ring that induces a *cis*-amide configuration and disrupts ß-sheet and α-helix formation. Psuedo-prolines are incorporated as dipeptide building blocks due to the poor reactivity and coupling efficiency of the secondary amine



Scheme 5.2: Total chemical synthesis of CCL2 *via* NCL. *i*: TMS-CHN_2, THF, Hexane. *ii*: Ethyl 3-mercaptopropionate, sodium thiophenolate, DMF. *iii*: TFA, 1,2-ethanedithiol, phenol, thioanisole, water, TIPS. *iv*: Gdn HCl, sodium phosphate, thiophenol, water.

All of the previous solid phase peptide syntheses of the chemokine CCL2 are summarised in **Table 5.2**.

Key Modification(s)	Strategy	Synthesizer	Coupling Reagents	Resin	Year	Purified* Yield	Author/Ref
Deletion of N-terminal residues	Stepwise Boc-SPPS	ABI; model 430A	DIC, DCC & HOBt	PAM (0.80 mmol/g)	1995	0.6-1.5%	Gong & Clark-Lewis.[19]
Ala substitution of Cys residues	Stepwise Boc-SPPS	ABI; model 430A	DIC, DCC & HOBt	PAM (0.80 mmol/g)	1997	0.6-1.5%	Gong & Clark-Lewis.[27]
N-Terminal Tbfmoc protection/probe	Stepwise Fmoc-SPPS	ABI; model 430A	Asymmetric anhydrides & HOBt	Wang (0.86 /mmolg)	1996	0.6%	Brown et al.[20]
Native CCL2 synthesis followed by global biotinylation	Stepwise Fmoc-SPPS	ABI: model 431A	HBTU & HOBt or DCC & HOBt	Low-loading Wang (0.12 mmol/gl)	2006	13-26%	Kruszynski et al.[23]
Site-specific on-resin biotinylation (Lys^{69} and Lys^{75})	Stepwise Fmoc-SPPS	ABI: model 431A	HBTU & HOBt or DCC & HOBt	Low-loading Wang (0.12 mmol/gl)	2006	18-20%	Kruszynski et al.[24]
Native CCL2 synthesis and co-crystallisation	Stepwise Fmoc-SPPS	ABI; model 433A	HBTU & HOBt	-	2006	-	Reid et al.[25]
Segmented synthesis *via* NCL and the use of structure-breaking pseudo-proline building blocks.	NCL and Fmoc-SPPS	ABI; model 431A	DCC & HOBt	Low-loading Wang (0.12 mmol/g) Sulfamylbutyryl (0.26 mmol/g)	2010	1.5% (with respect to thioester fragment)	Grygiel et al.[9]

Table 5.2: Previous room temperature solid phase peptide syntheses of CCL2. *>95% purity by RP-HPLC.

5.2.4 Preliminary Work on the Microwave Assisted Linear Synthesis of CCL2

* *The preliminary work reported here is a summary of the investigation carried out by Mr. S. Lear an MRes student in the Cobb group.*

CCL2 has been shown to be amenable to a complete stepwise linear peptide synthesis with procedures for the purification and subsequent folding into the correct active protein also known.[9,20] In the previously described syntheses low-loading resins, double couplings and pseudo-prolines were all utilised to enable pure product CCL2 to be obtained.

It was envisaged an improved (particularly in terms of synthesis time) linear stepwise synthesis of CCL2 that utilised microwave assisted SPPS and a low-loading resin.

Microwave irradiation/heating can significantly reduce synthesis times and improve crude yields. The precise nature of heating[28**] that can be applied has shown to be of specific benefit when applied to sterically hindered couplings and the synthesis of sequences that are prone to secondary structure formation/aggregation.[29] It was postulated that microwaves provided an additional heating of peptides directly through interactions with amide bond dipoles which was in part responsible for the elevated yields over room temperature SPPS.[30-32] However, work by the Kappe group comparing conventional (oil bath) and microwave heating in the synthesis of three difficult peptides provided evidence against this hypothesis. The syntheses showed near identical improvements in peptide product purity and similar (low) racemisation levels showing that the main effect of microwave heating (in relation to SPPS) is purely thermal.[33]

The majority of the previous syntheses of CCL2 used a low-loading Wang resin. Therefore, a polystyrene low-loading Wang (PABA) resin (0.27 mmol/g) was chosen. Low loading resins are recommended for the synthesis of long or difficult sequences as they reduce the probability for disadvantageous steric interactions between peptide chains that can confine reactive sites and truncate the synthesis.[34,35]

The resin was loaded onto a CEM liberty1 automated synthesizer and N-Fmoc amino acids with their appropriate side-chain protecting group were added in a stepwise

[**]Microwave heating disposes energy to the solvent and reagents directly and their storage and conversion of the energy to heat varies. DMF, the common SPPS solvent has good microwave heating ability (loss factor: tan δ 0.161).

fashion until the desired polypeptide sequence was formed. The coupling conditions that were used are shown in **Table 5.3**.

Reaction	Note(s)	Reagents & Equivalents	Temp.	Time
General peptide coupling	Microwave assisted	PyBOP (5.0 eq), DIPEA (10 eq), Fmoc-AA-OH (5.0 eq).	75 °C	10 min
Specific peptide coupling (His)	Microwave assisted	PyBOP (5.0 eq), DIPEA(10 eq), Fmoc-His-OH (5.0 eq).	50 °C	10 min
Fmoc deprotection	Microwave assisted	Piperdine (20 %), DMF (80%).	75 °C	3 min
TFA cleavage	-	TFA (95%), TIPS (2.5%), H_2O (2.5%).	RT	3 h

Table 5.3: Reaction details for the all SPPS steps used in this synthesis.

The synthesis was monitored with a series of test-cleavages[††] and it showed that the synthesis failed between amino acids 35 – 52 (**Figure 5.4**). This region contains hydrophobic residues thus, has a high aggregation potential. In the native protein this a section of the sequence folds into β-sheets.

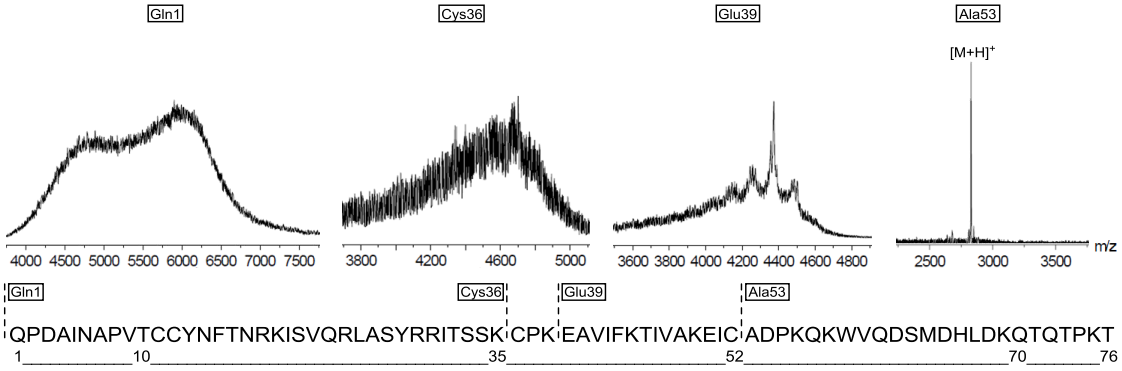


Figure 5.4: *Top*: The MALDI-TOF spectra of crude peptide isolated from test cleavages at four locations. *Bottom*: The native sequence of CCL2 with the relevant test-cleave locations indicated.

[††]A test cleavage enables the progress of a solid phase synthesis to be monitored and purity of product peptide to be tentatively described. A very small amount of resin can be removed and treated with TFA cleavage "cocktail" for a set time (generally ~ 3 h). Subsequent evaporation of the TFA and ether washes yields a small quantity of crude peptide that can be analysed *via* MS and RP-HPLC. Test cleavages are also particularly useful in the optimisation of cleavage conditions.

The previous work guided the progress of a microwave assisted synthesis of CCL2 to utilise either a fragmentation or an optimised stepwise approach:

1. *A fragmentation approach*: To divide CCL2 into smaller peptide segments that would trivialise their synthesis and form the full native CCL2 using NCL reactions as in the Grygiel *et al*. synthesis. In this approach the synthesis of a peptide thioester C-terminus is required, to do this a selection of "activated" resins have become commercially available and a set of known post-synthesis steps yield the peptide thioester.[9]

2. *An optimised stepwise approach*: The optimisation would involve resins, reaction times, reaction temperatures and amino acids so that a new stepwise linear synthesis could progress through the aforementioned region of hydrophobic amino acids and form the full native CCL2.

5.3 Microwave Assisted Total Chemical Synthesis of CCL2: A Fragmentation Approach

5.3.1 A Fragment Approach to the Total Synthesis of CCL2

CCL2 contains 4 cystiene residues and these residues allow various segmented approaches to be hypothesised (two routes are shown in **Table 5.4** below). This synthetic approach originated from the failure of the synthesis after ~ 24 coupling steps. Hence, all the hypothesised routes divide the sequence at Cys52 to form fragment **52-76**. Each route contains at least one NCL (*i* in **Table 5.4**) and the linear stepwise synthesis of at least two fragments. Therefore, an "activated" resin is required to synthesise the C-terminal thioester or thioester "mimic" segment.

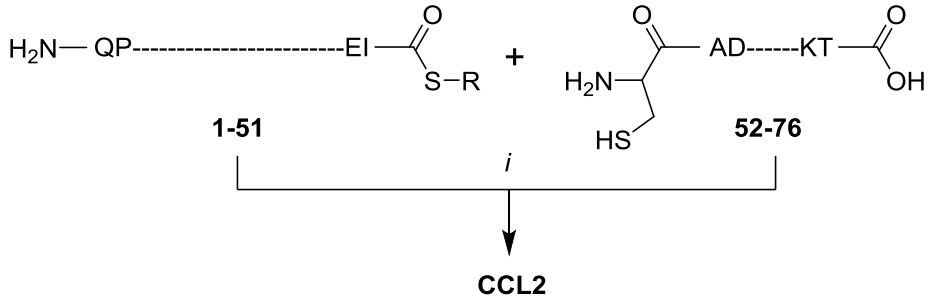
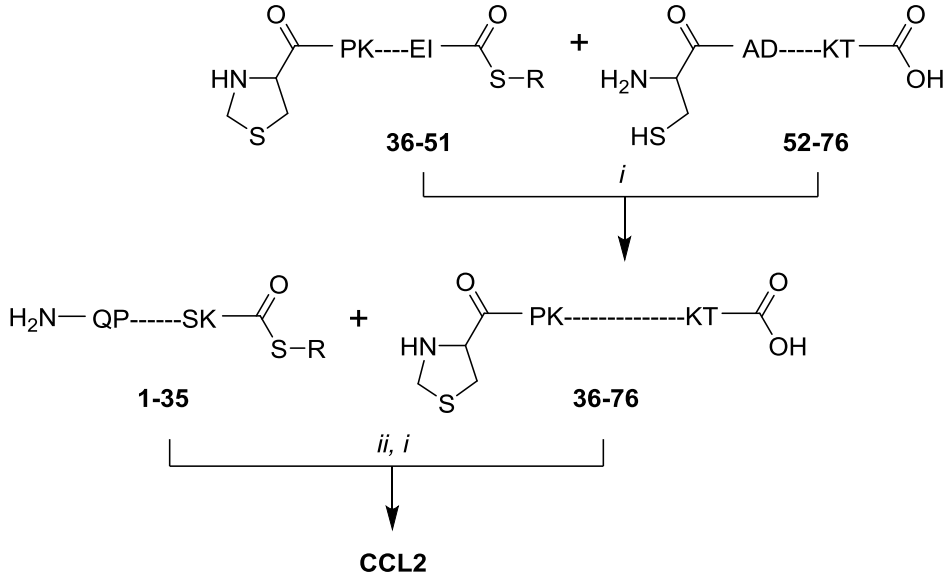
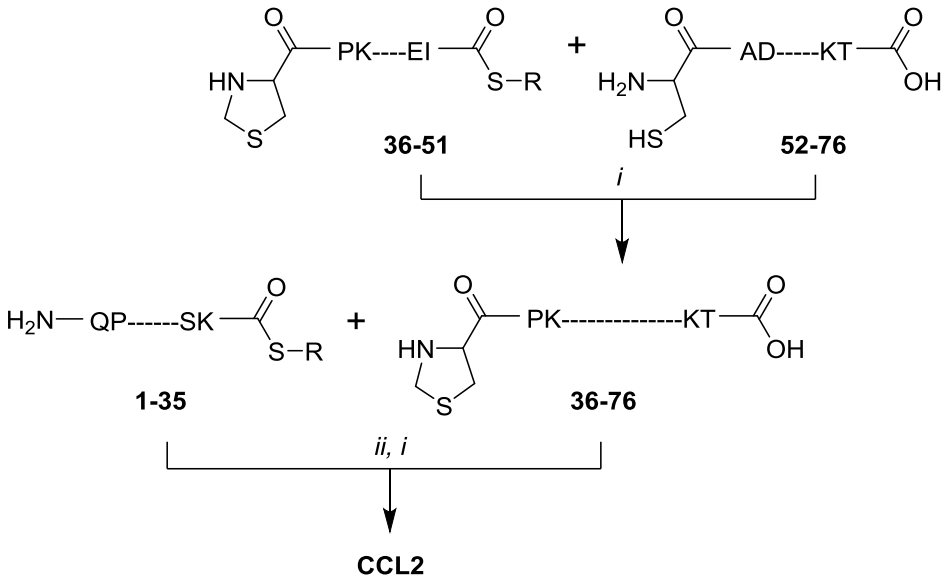
Route	Fragment Approach
1	
2	

Table 5.4: Proposed routes to the synthesis of chemokine CCL2. *Route 1*: fragments **1-51** and **52-76**. *Route 2*: fragments **1-35**, **36-51** and **52-76** approach. *i*: Native chemical ligation (NCL). *ii:* Deprotection of thiazolidine N-terminal cysteine protection.

We focused on the recently developed Dawson Dbz resin for the SPPS of the required thioesters. Dawson Dbz was implicated as a resin that would show improvements on the low yields and lack of facile (test-cleavage) reaction monitoring of the historically used sulfamylbutyryl resin.[14] Preliminary work by S. Lear on a small peptide showed that the microwave heating conditions could not provide the desired peptide product as a single species. The impurities that arose during the synthesis were tentatively assigned as multiple amino acid additions propogating from the alternative aromatic amine. However, Alloc protection of the non-propagating aromatic amine solved this issue in a similar way to glycine rich peptides.[36]

There is precidence for microwave assisted synthesis on a Dawson Dbz resin[17] but, the synthesis was optimised and microwave heating was not used for all couplings e.g. the first three amino acids were coupled at room temperature with specific and individual coupling reagents, reagent equivalents, reaction times and repetitions. Therefore, our fragmented approach will utilise an Alloc protected Dawson Dbz resin in a complete microwave assisted synthesis of the thioester/thioester "mimic" fragments.

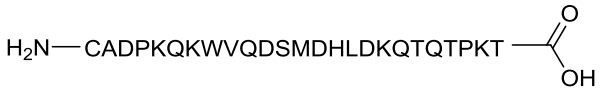
The aforementioned Ile-Cys junction that is vital in all possible strategies has been shown previously to be extremely poor for NCL. **Route 1** requires the synthesis of the longer peptide: **1-51** that could potentially be more challenging and time-consuming. Whereas, in **Route 2**, the synthesis of the shorter fragment **36-51** and the subsequent NCL with pre-synthesised fragment **52-76** would enable valuable knowledge of the use of thioester/"mimic" resins as well as ligation at the Ile-Cys junction in a shorter time-frame. Therefore, **Route 2** (**Table 5.4** and **Scheme 5.3**) was chosen.



Scheme 5.3: Synthesis of CCL2 from three fragments: **1-35**, **36-51** and **52-76**: *i*: NCL. *ii:* Deprotection of thiazolidine N-terminal cysteine protection .

5.3.1.1 Synthesis of CCL2 Peptide Fragment 52-76

The microwave assisted Fmoc SPPS of CCL2 peptide fragment **52-76** (**Figure 5.5**) is previously described in **Section 5.2.4**.



52-76
2928.4 Da

Figure 5.5: Primary sequence of fragment **52-76** from N-terminal amine and C-terminal carboxylic acid

The re-synthesis was carried out as in **Section 5.2.4**, **Table 5.3**. However, PyBOP and DIPEA were replaced by the more commonly used in microwave assisted synthesis coupling reagents: DIC and HOBt see **Table 5.5**. DIC/HOBt (or similar) has been effectively used to synthesize (with microwave heating) many difficult sequences.[37-40]

Reaction	Note(s)	Reagents & Equivalents	Temp.	Time
General peptide coupling	Microwave assisted	DIC (5.0 eq), HOBt (10 eq), Fmoc-AA-OH (5.0 eq).	75 °C	10 min
Specific peptide coupling (His)	Microwave assisted	DIC (5.0 eq), HOBt (10 eq), Fmoc-His-OH (5.0 eq).	50 °C	10 min
Fmoc deprotection	Microwave assisted	Piperdine (20 %), DMF (80%).	75 °C	3 min
TFA cleavage	-	TFA (95%), TIPS (2.5%), H_2O (2.5%).	RT	3 h

Table 5.5: Reaction details for the all SPPS steps used in this synthesis.

The expected mass of CCL2 peptide fragment **52-76** was observed *via* MALDI-TOF analysis with a $[M+H]^+$ peak being observed at 2928.8 Da. Crude peptide was obtained at a purity of 80% and was purified *via* preparative RP-HPLC. The final peptide was obtained in 17% yield and shown to have a purity of greater than 95% (**Figure 5.6**).

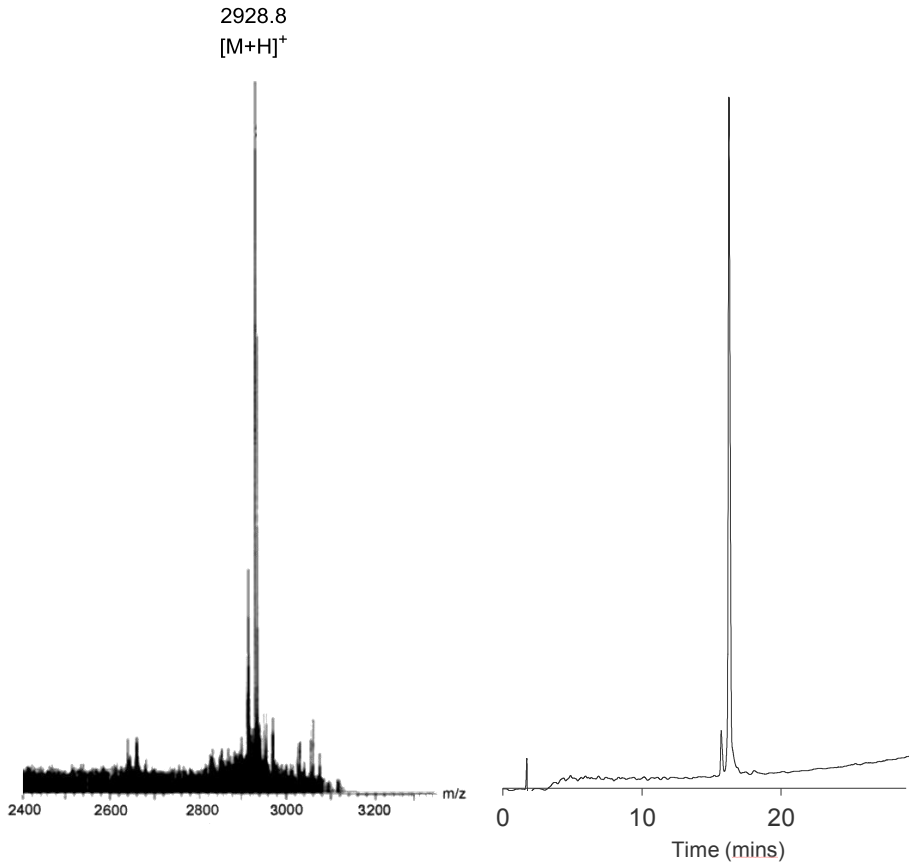
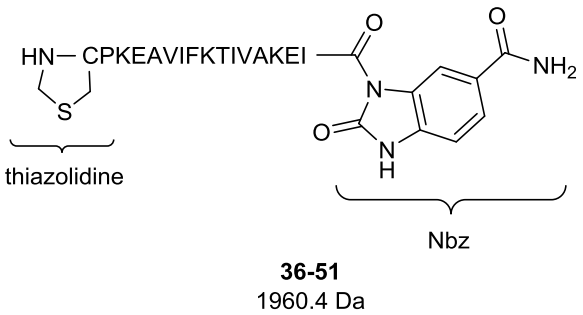


Figure 5.6: Analysis of the purified CCL2 peptide fragment **52-76**. *Left*: positive ion MALDI-TOF spectrum. *Right*: Analytical RP-HPLC (230 nm). The linear gradient ranges from 2% MeCN (0.1% TFA) in H_2O (0.1% TFA) at 0 minutes to 50% MeCN (0.1% TFA) in H_2O (0.1% TFA) at 30 minutes.

5.3.1.2 Synthesis of CCL2 Peptide Fragment 36-51

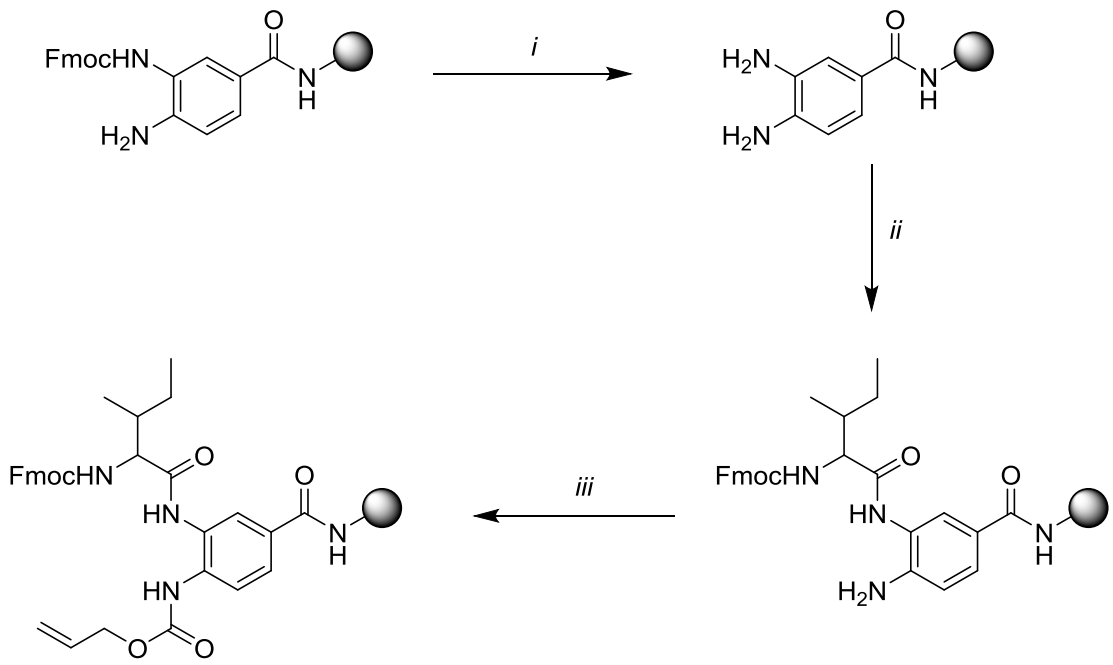
Commercially available low-loading (PEG/PS hybrid) *N*-Fmoc Dawson Dbz was used to prepare the C-terminal *N*-acylurea (thioester mimic) Nbz peptide (**Figure 5.7**).



36-51
1960.4 Da

Figure 5.7: CCL2 peptide fragment **36-51**: containing thiazolidine protected N-terminal Cysteine and *N*-acylurea (Nbz) C-terminus.

The first synthetic step in the preparation of peptide fragment **36-51** is loading of the Dbz resin with the first amino acid building block: Fmoc-Ile-OH. The standard[36] resin loading conditions (including; non-propagating aromatic amine Alloc protection) are shown in **Scheme 5.4**.



Scheme 5.4: Loading the Dbz resin. *i*: Piperidine/DMF (1:4), RT, 10 min, x 2. *ii*: PyBOP (5 eq), Fmoc Ile-OH (5 eq), DIPEA (10 eq), DMF, RT, 1 h. *iii*: DIPEA (1.1 eq) allyl chloroformate (350 nM), DCM, RT, 1 h.

From the loaded Dbz resin the synthesis of CCL2 peptide fragment **36-51** was undertaken on a CEM liberty1 automated synthesizer (**Table 5.6**). The post synthesis *N*-Alloc deprotection and Dbz cyclisation that lead to the final Nbz containing CCL2 peptide fragment **36-51** were conducted manually at room temperature (**Table 5.6**).

Reaction	Note(s)	Reagents & Equivalents	Temp.	Time
General peptide coupling	Microwave assisted coupling repeated	DIC (5.0 eq), HOBt (10 eq), Fmoc-AA-OH (5.0 eq).	75 °C	10 min
Specific peptide coupling (Boc-thiazolidine-OH)	Coupling repeated until negative TNBS/Kaiser resin test result.	PyBOP (2.5eq), DIPEA (5.0 eq), Boc-thiazolidine-OH (2.5 eq).	RT	1 h 30 min
Fmoc deprotection	Microwave assisted	Piperdine (20 %), DMF (80%).	75 °C	3 min
Alloc deprotection	Repeated	Tetrakis(triphenylphosphine) palladium$^{(0)}$ (0.35 eq), phenylsilane (20 eq), DCM.	RT	30 min
Dbz cyclisation (Nbz formation)	Repeated	1.4-nitrophenylch-loroformate (5.0 eq.), DCM. 2. DIPEA, DMF.	RT	1. 1 h 2. 45 min
TFA cleavage	-	TFA (95%), TIPS (2.5%), H_2O (2.5%).	RT	3 h

Table 5.6: Reaction details for the all SPPS steps used in this synthesis.

The expected mass for CCL2 Peptide Fragment **36-51** was observed *via* MALDI-TOF analysis with a $[M+H]^+$ peak being observed at 1960.4 Da. Crude CCL2 Peptide Fragment 36-51 was obtained at a purity of 32.7% and subsequent purification of the crude peptide was achieved *via* preparative RP-HPLC. The final peptide was shown to have a purity of 94% (**Figure 5.8**) and it was obtained in 14% yield. The MALDI-TOF spectrum shows a clear impurity peak at 1800.4 Da that is attributed to the hydrolysis of the *N*-acylurea to yield the C-terminal carboxylic acid. This impurity was also visualised in the ESI+ MS (denoted * in **Figure 5.8**). However, due to the sharp analytical RP-HPLC trace it was assumed that fragment **31-51** was prepared as a single species and that the ionisation conditions in both MS techniques led to the observed hydrolysis.

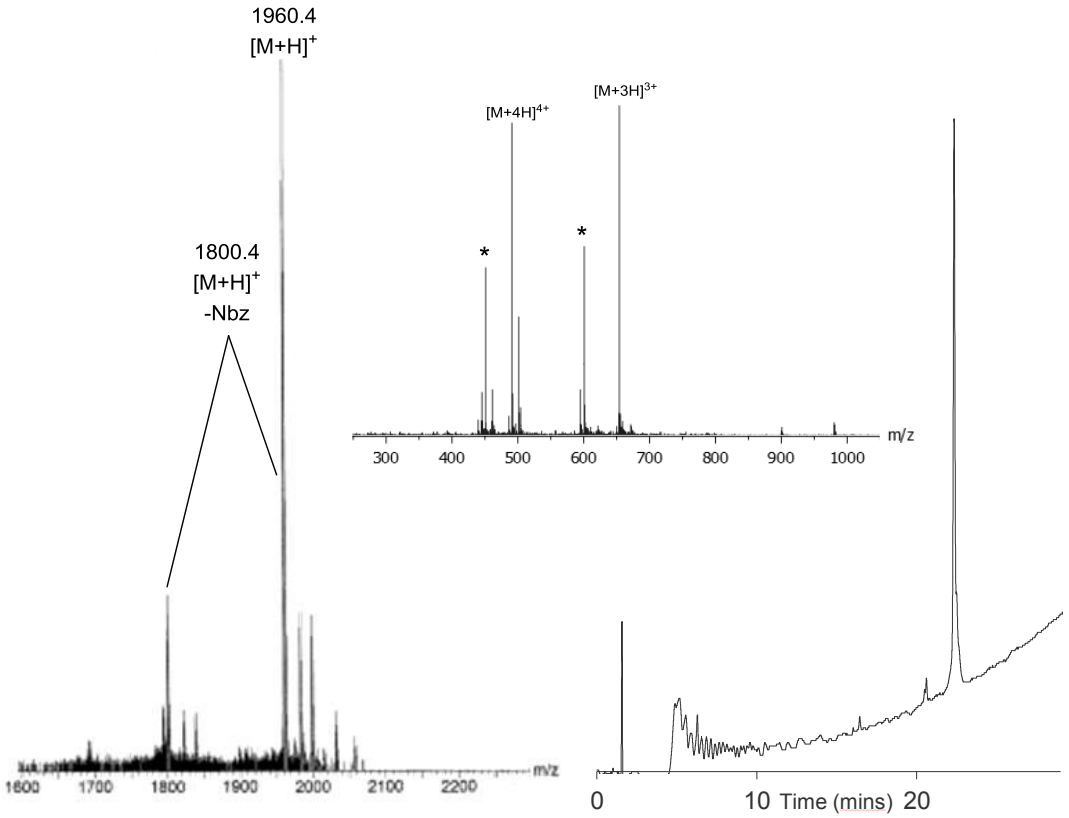
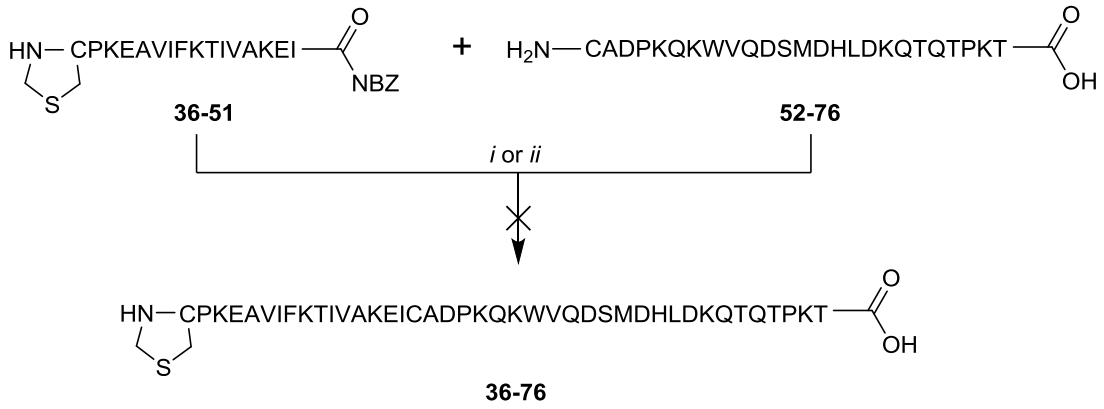


Figure 5.8: Analysis of the purified CCL2 fragment **31-51**. *Left*: positive ion MALDI-TOF spectrum. *Middle*: ESI+ spectrum. *Right*: Analytical RP-HPLC (230 nm) of Fragment 31-51. The linear gradient ranges from 2% MeCN (0.1% TFA) in H_2O (0.1% TFA) at 0 minutes to 50% MeCN (0.1% TFA) in H_2O (0.1% TFA) at 30 minutes.

5.3.1.3 NCL of CCL2 Peptide Fragments 36-51 and 52-76

With two CCL2 peptide fragments in hand (**36-51** and **52-76**), the NCL reaction was carried out based on the NovaBiochem procedure[14] on a small test scale (3 mg) (*i* in **Scheme 5.5**). The procedure was repeated without guanadinium hydrochloride (Gdn.HCl) (*ii* in **Scheme 5.5**). In either case no trace of ligation product was observed at 4, 8, or 24 h in MALDI-TOF or LCMS spectra. After 24 h The N-acylurea was seen to undergo complete hydrolysis into the carboxylic acid (1823.5 Da, $[M+Na]^+$-Nbz, **Figure 5.9**).



Scheme 5.5: Attempted NCL to form CCL2 peptide fragment **36-76**. *i:* MPAA, sodium phosphate, TCEP, Gdn HCl, H_2O, RT, 18 h. *ii:* MPAA, sodium phosphate, TCEP, H_2O, RT, 24 h.

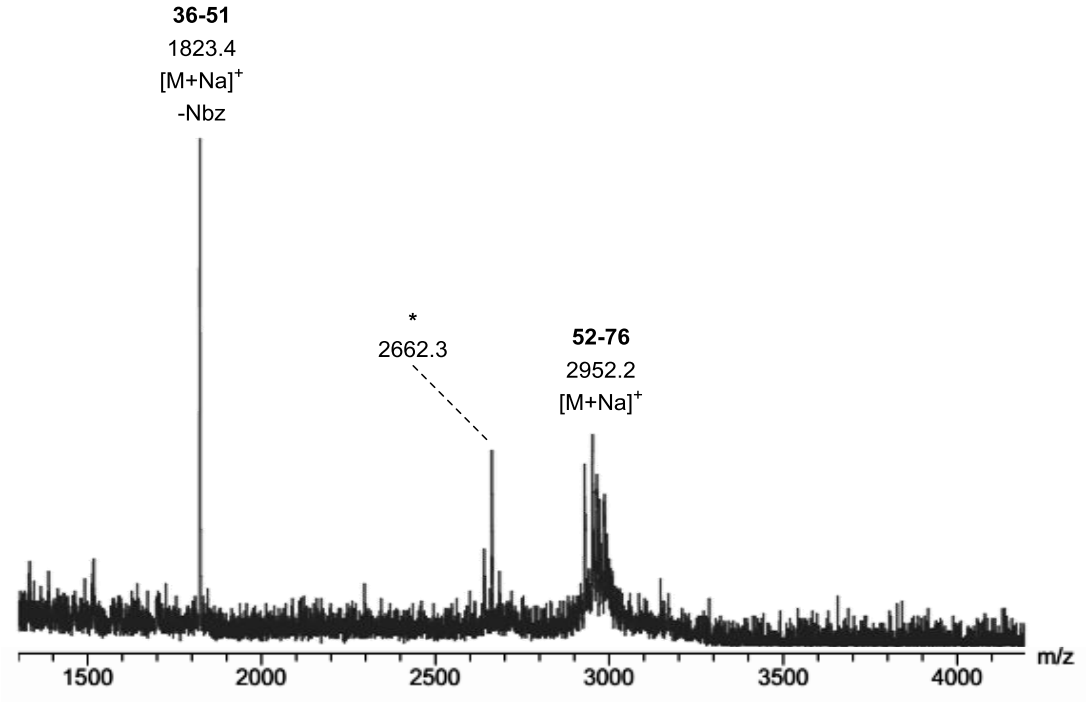


Figure 5.9: MALDI-TOF spectrum of the crude NCL mixture after 24 h. *denotes unassignable peak.

It was deemed that the NCL reaction at the Ile51-Cys52 junction was going to be slow and thus, we focused on an optimised linear SPPS approach that would enable a successful synthesis beyond this junction. The hope was that by optimising the SPPS procedures (in **Section 5.2.4** and **Section 5.3.1.1**) that, a full linear synthesis of CCL2 or a synthesis up to the next possible ligation point at Lys35-Cys36.

5.4 Microwave Assisted Total Chemical Synthesis of CCL2: An Optimised Stepwise Approach

5.4.1 An Optimised Stepwise Approach to the Total Synthesis of CCL2

We envisaged an "improved" strategy to enable the synthesis to progress past point at which the previous failed stepwise synthesis failed: Cys52 (**Section 5.1.4**). To do this a number of modifications to the original synthesis were made:

1. Double (microwave assisted) couplings throughout. (*Standard Conditions*, **Figure 5.10**)
2. Increased temperature couplings in the parts of the sequence with significant "aggregation" potential (where the original synthesis failed). (*High Temperature Conditions*, **Figure 5.10**)
3. Replace the low-loading polystyrene based Wang resin with a low-loading hybrid PEG/PS tentagel resin that has increased swelling properties (leading to a reduction in the effects of aggregation).[41]
4. Monitor the previously difficult stages of the synthesis (via test-cleavage) and utilise pseudo-proline amino acids if the synthesis was judged to be poor.

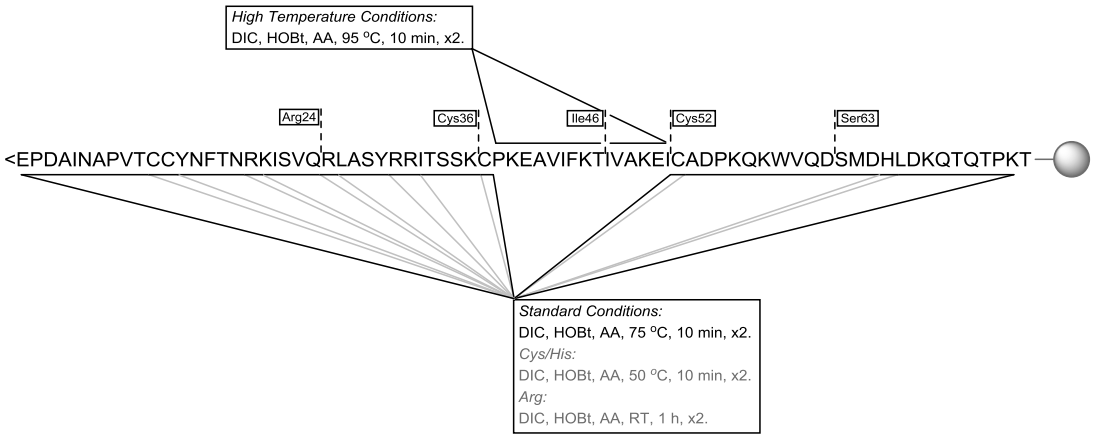


Figure 5.10: Breakdown of proposed reaction conditions for optimised SPPS of CCL2. Test cleavage points are indicated at Ser 63, Cys52, Ile46, Cys36 and Arg24.

To begin, pre-loaded (Thr(O*t*Bu)) Tentagel resin was loaded onto a CEM liberty1 automated synthesizer and SPPS was undertaken with reagents and conditions as in **Table 5.7**.

Reaction	Note(s)	Reagents & Equivalents	Temp.	Time
General peptide coupling	Microwave assisted coupling repeated	DIC (5.0 eq), HOBt (10 eq), Fmoc-AA-OH (5.0 eq)	75 °C	10 min
High temperature peptide coupling (**PKEAVI** & **IVAKEI**)	Microwave assisted coupling repeated	DIC (5.0 eq), HOBt (10 eq), Fmoc-AA-OH (5.0 eq).	95 °C	10 min
Specific peptide coupling (Cys & His)	Microwave assisted coupling repeated	DIC (5.0 eq), HOBt (10 eq), Fmoc-AA-OH (5.0 eq).	50 °C	10 min
Specific peptide coupling (pseudo-proline (*KT*))	Coupling repeated until negative chloranil resin test result.	PyBOP (2.5 eq), DIPEA (5.0 eq), Fmoc- Lys(Boc)-Thr(Ψpro)-OH (2.5 eq).	RT	1 h 30 min
Fmoc deprotection	Microwave assisted	Piperdine (20%), DMF (80%).	75 °C	3 min
TFA cleavage	-	TFA (95%), TIPS (2.5%) H_2O (2.5%).	RT	3 h
Reductive TFA cleavage	-	*1*. TFA (95%), TIPS (5%). *2*. add: EDT, TMSBr	RT	*1*. 2 h 30 min. *2*. 30 min

Table 5.7: Reaction details for the all SPPS steps used in this synthesis.

Test cleavages were taken to monitor the synthesis as it progressed. This first test cleavage was carried out at Ser63 (**Figure 5.11**) and the test cleaved peptide showed a good level of purity by analytical RP-HPLC (**Figure 5.12**). In addition the major peaks observed in the MALDI-TOF spectrum corresponded to $M+H^+$ (1628.8 Da) and $M+Na^+$ (1659.8 Da) for the peptide fragment.

63-76
1628.8 Da

Figure 5.11: Chemical structure of CCL2 peptide fragment **63-76**.

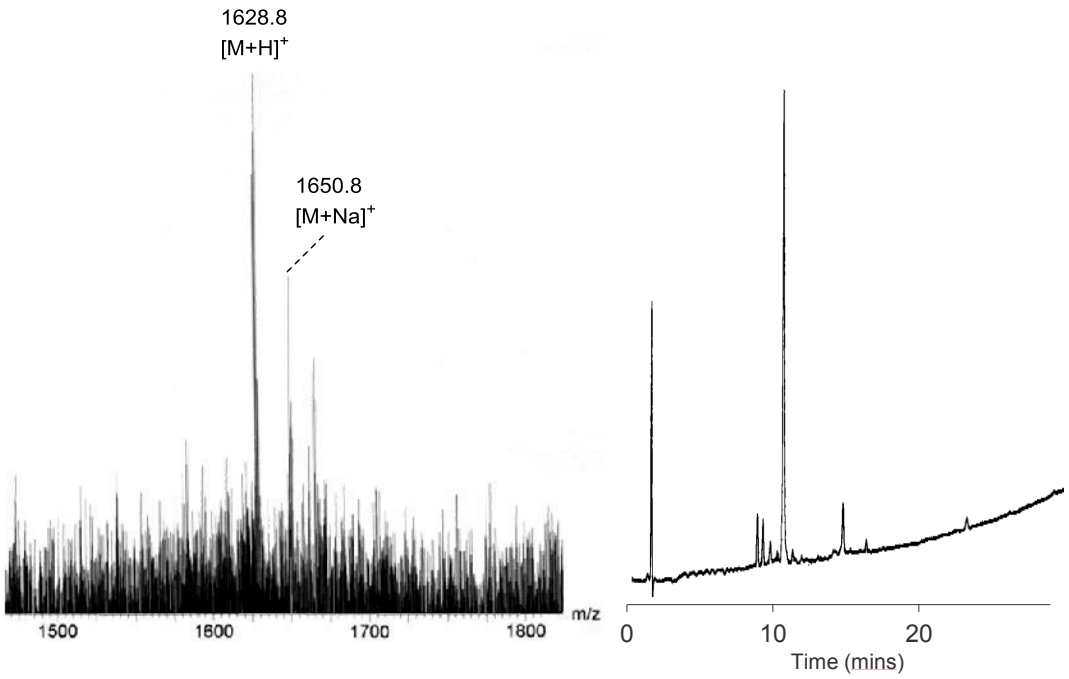
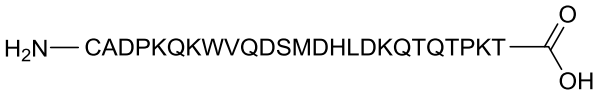


Figure 5.12: Analysis of the crude CCL2 peptide fragment **63-76**. *Left*: positive ion MALDI-TOF spectrum. *Right*: Analytical RP-HPLC (230 nm) of Fragment 52-76. The linear gradient ranges from 2% MeCN (0.1% TFA) in H_2O (0.1% TFA) at 0 minutes to 50% MeCN (0.1% TFA) in H_2O (0.1% TFA) at 30 minutes. The relatively low signal to noise ratio in the Maldi spectrum (positive ionisation mode) was attributed to the large number of residues that "prefer" a negative charged state (Ser, Thr and Asp) in this point in the sequence.

The synthesis was continued up to residue Cys52 and another test cleavage was undertaken **Figure 5.13**. The test cleaved peptide showed two distinct peptide peaks of similar intensities in the analytical RP-HPLC and LCMS-ESI^+ analysis confirmed that there was a 16 Da difference between the two main species (**Figure 5.14**). Initial observations deduced that this was due to methionine oxidation and that a reduction was needed. The oxidation of methionine was not observed in the previous stepwise synthesis (**Section 5.1.4**). Therefore, the prolonged reaction times (due to double coupling) and/or higher swelling Tentagel resin must facilitate the oxidation. An optimised TFA cleavage with the addition of EDT and TMSBr (reductive TFA cleavage conditions, **Table 5.7**) fully and selectively reduced the methionine, hence alleviating the problem.



52-76
2927.4 Da

Figure 5.13: Chemical structure of CCL2 peptide fragment **52-76**.

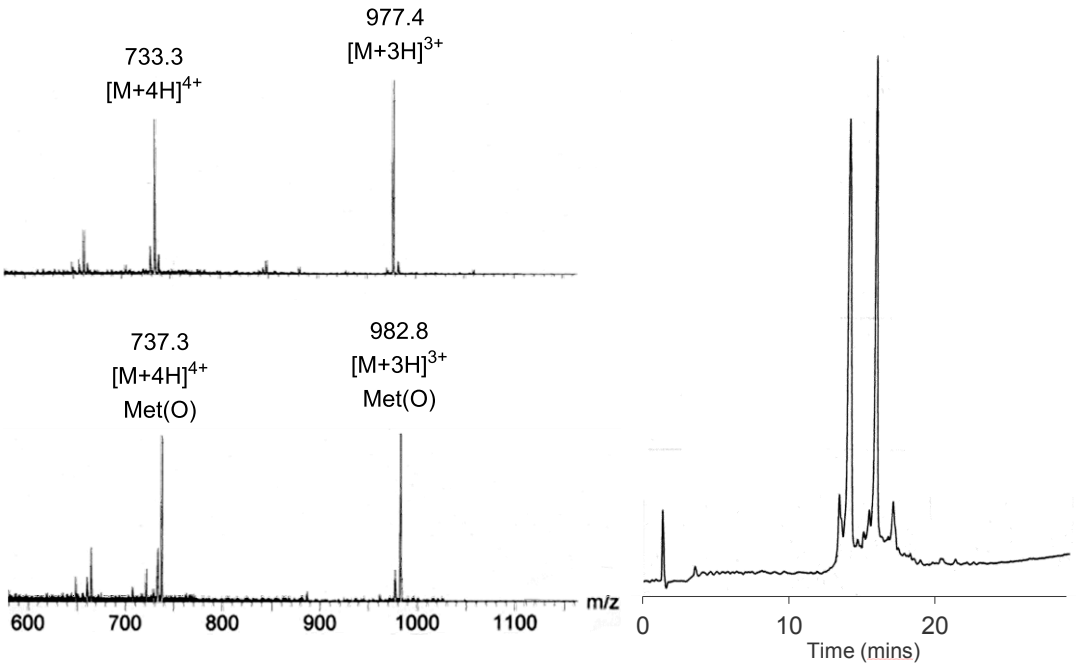


Figure 5.14: Analysis of the crude CCL2 peptide fragment **52-76**. *Left bottom*: LCMS-ESI$^+$ spectrum of front running major peak. *Left top*: LCMS-ESI$^+$ spectrum of late running major peak. *Right*: Analytical RP-HPLC (230 nm). The linear gradient ranges from 2% MeCN (0.1% TFA) in H_2O (0.1% TFA) at 0 minutes to 50% MeCN (0.1% TFA) in H_2O (0.1% TFA) at 30 minutes.

As the Synthesis progressed into the hydrophobic "prone to aggregation" region of CCL2, higher temperature (95°C) couplings (**Table 5.7**) were used up to Ile46 (**Figure 5.15)**. Test cleavage (with reducing conditions to avoid Met oxidation) showed a major peak corresponding to the desired peptide (by LCMS, **Figure 5.16**). However, distinct yet unassignable impurity peaks in the LCMS (* in **Figure 5.16**) were visible. In addition the major peak that corresponded to the target peptide showed broadening in the analytical RP-HPLC. The peak broadening in the RP-HPLC spectrum LCMS was thought to be due to aggregation/peptide folding leading to less effective peptide coupling steps. Therefore, the pseudo-proline KT was incorporated into the sequence in an attempt to "break-folding" and aid the efficiency of the subsequent coupling steps.

46-76
3580.8 Da

Figure 5.15: Chemical structure of CCL2 peptide fragment **46-76**.

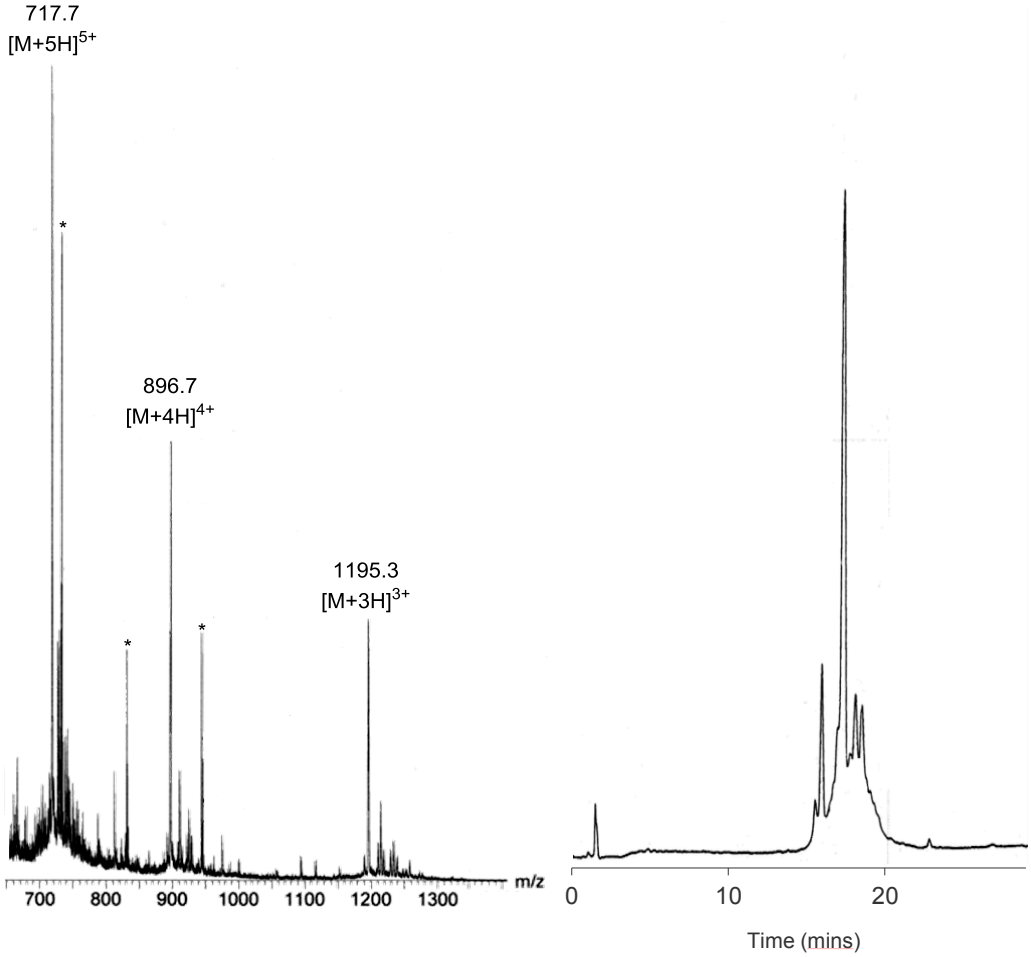
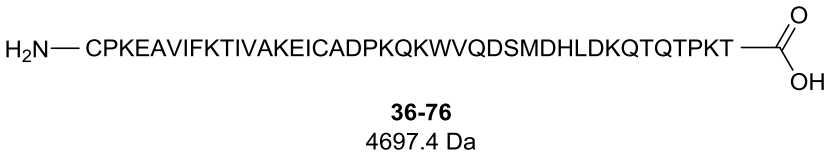


Figure 5.16: Analysis of the crude CCL2 peptide fragment **46-76**. *Left*: ESI+ spectrum. *Right*: Analytical RP-HPLC (230 nm). The linear gradient ranges from 2% MeCN (0.1% TFA) in H_2O (0.1% TFA) at 0 minutes to 50% MeCN (0.1% TFA) in H_2O (0.1% TFA) at 30 minutes. *unassignable impurity peaks.

After manual coupling of KT pseudo-proline dipeptide the synthesis was continued using higher temperature (95°C) couplings (**Table 5.7**) up to Cys36 (**Figure 5.17**) as this is the next possible ligation junction. The test cleaved peptide showed multiple peaks of very similar retention time in the analytical RP-HPLC and the LCMS of crude peptide showed desired product peaks ($[M+5H]^{5+}$, $[M+6H]^{6+}$, $[M+7H]^{7+}$ and

$[M+8H]^{8+}$) as well as a significant amount of non-product peaks (**Figure 5.18**). It was not possible to assign the non-product peaks as particular species.



36-76
4697.4 Da

Figure 5.17: Chemical structure of CCL2 peptide fragment **36-76**.

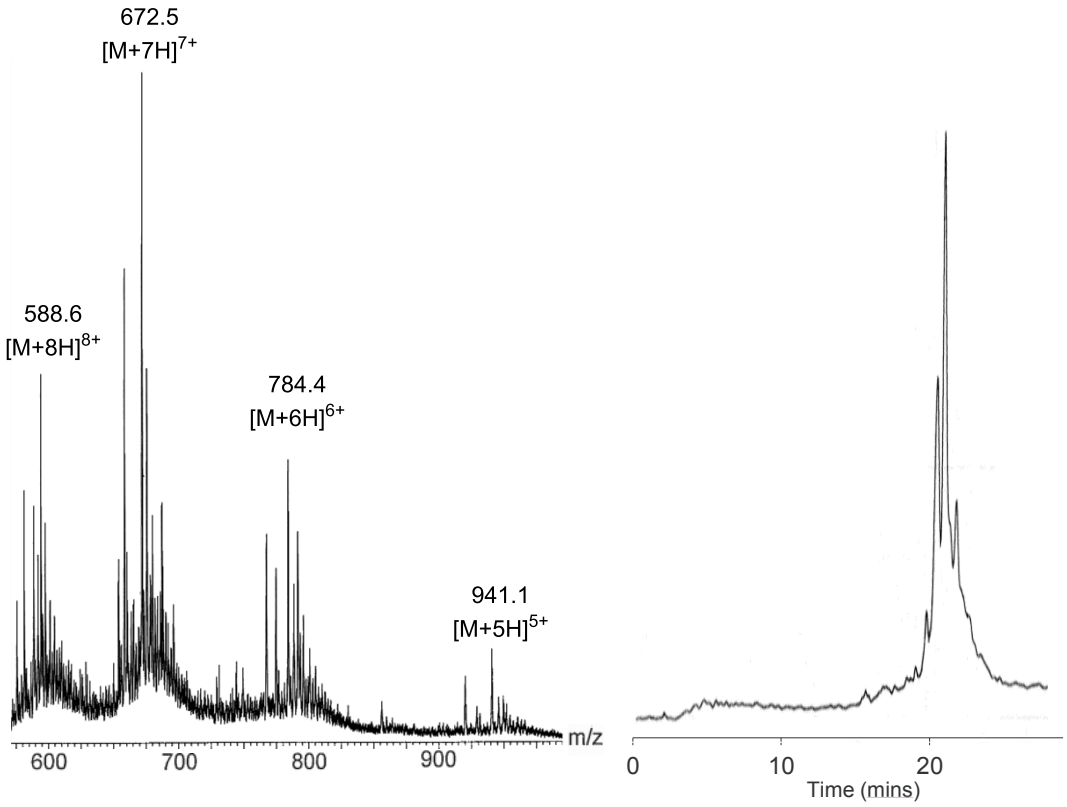


Figure 5.18: Analysis of the crude CCL2 peptide fragment **36-76**. *Left*: ESI+ spectrum. *Right*: Analytical RP-HPLC (230 nm). The linear gradient ranges from 2% MeCN (0.1% TFA) in H_2O (0.1% TFA) at 0 minutes to 50% MeCN (0.1% TFA) in H_2O (0.1% TFA) at 30 minutes.

Although the product peptide (residue **36-76**) is believed to be the major product, it was felt that if the synthesis was continued then isolation of a single species *via* semi-preparative RP-HPLC would not be trivial. In addition, if the synthesis was continued the addition of more impurities would compound this problem. Therefore, the resin was cleaved yielding crude CCL2 Peptide Fragment **36-51** at a purity of 32.7%. Purification *via* RP-HPLC afforded CCL2 Peptide Fragment **36-51** in 8%

yield at a purity of 96% (**Figure 5.19**). The synthesis of this 41 amino acid fragment is a significant improvement on the previous synthetic strategy (**Section 5.2.4**) that was not able to produce product peptide after amino acid Cys52 (a 25 amino acid fragment).

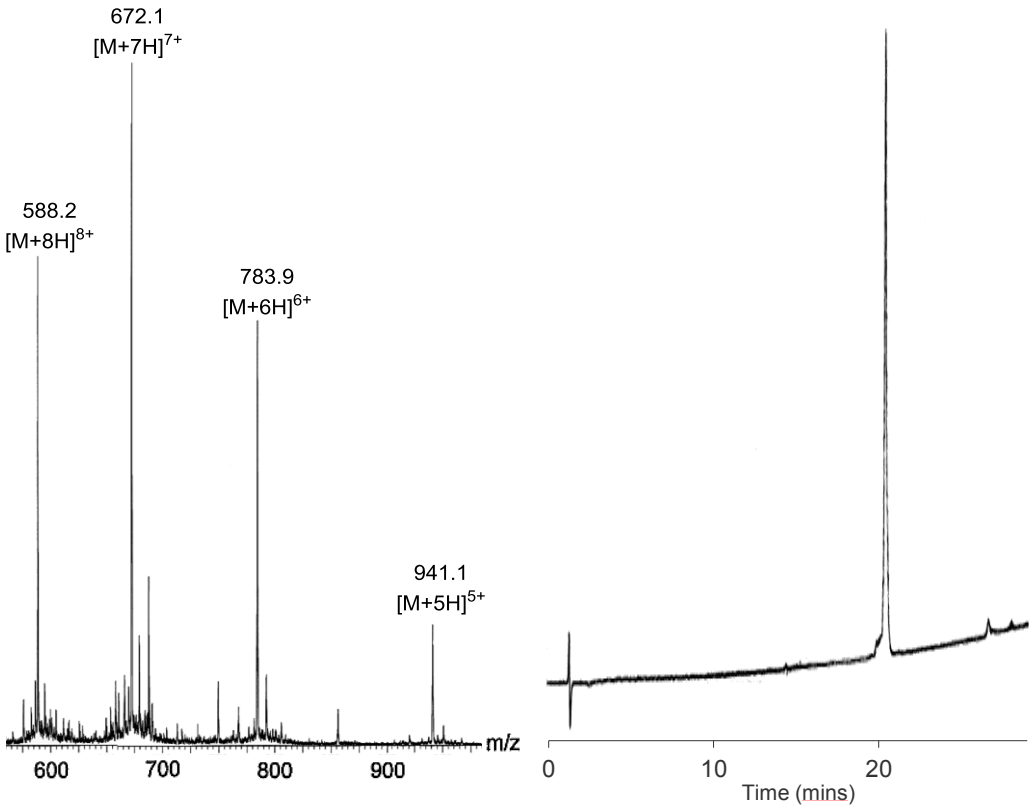


Figure 5.19: Analysis of the purified CCL2 peptide fragment **36-76**. *Left*: ESI+ spectrum. *Right*: Analytical RP-HPLC. The linear gradient ranges from 2% MeCN (0.1% TFA) in H_2O (0.1% TFA) at 0 minutes to 50% MeCN (0.1% TFA) in H_2O (0.1% TFA) at 30 minutes.

The cleavage and purification of CCL2 peptide fragment **36-76** means that a fragmented approach must be used and that the N-terminal section of CCL2 peptide fragment **1-35** must be synthesised with a C-terminal thioester or thioester "mimic": As seen in **Section 5.3.1.2**, a protected Dawson Dbz strategy was chosen. This leads to a 2-segment strategy similar to the Grygiel *et al.* synthesis.[9]

5.4.2 Synthesis of CCL2 Peptide Fragment 1-35

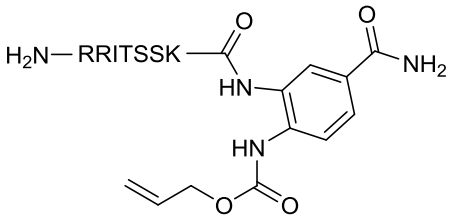
With CCL2 peptide fragment **36-76** in hand; synthesis of CCL2 peptide fragment **1-35** was undertaken on a Dawson Dbz resin. Firstly, the Dbz resin was loaded with Fmoc-Lys(Boc)-OH and protected with allyl chloroformate (as in **Scheme 5.4** in **Section 5.3.1.2**). The pre-loaded and protected resin was loaded onto a CEM liberty1 automated synthesizer and the synthesis was attempted (for conditions see: **Table 5.8**). As with the previous synthesis (peptide fragment **36-76**) test cleavages monitoring the synthesis were taken at regular intervals.

Reaction	Note(s)	Reagents & Equivalents	Temp.	Time
General Peptide coupling	Repeated	DIC (5.0 eq), HOBt (10 eq), Fmoc-AA-OH (5.0 eq)	75 °C	10 min
Peptide coupling (Cys)	Repeated	DIC (5.0 eq), HOBt (10 eq), Fmoc-AA-OH (5.0 eq).	50 °C	10 min
Peptide coupling (Arg)	Repeated	DIC (5.0 eq), HOBt (10 eq), Fmoc-AA-OH (5.0 eq).	RT	1 h
Peptide coupling (Boc-pyroglutamic acid)	Repeated until negative chloranil resin test result.	PyBOP (2.5 eq), DIPEA(5 eq), Fmoc-Lys(Boc)-Thr(Ψpro)-OH (2.5 eq).	RT	1 h 30 min
Fmoc deprotection	-	Piperdine (20 %), DMF (80%).	75 °C	3 min
Alloc deprotection	Repeated	Tetrakis(triphenylphosphine) palladium$^{(0)}$ (0.35 eq), phenylsilane (20 eq), DCM.	RT	30 min
Dbz cyclisation (Nbz formation)	Repeated	1.4-nitrophenylch-loroformate (5.0 eq.), DCM. 2. DIPEA, DMF.	RT	1. 1 h 2. 45 min
TFA cleavage	-	TFA (95%), TIPS (2.5%) H_2O (2.5%).	RT	3 h

Table 5.8: Reaction details for the all SPPS steps used in this synthesis.

The first test cleavage of the peptide was at Arg29 (**Figure 5.20**) and RP-HPLC analysis showed that the crude peptide had a good level of purity (**Figure 5.21**). In

addition, $[M+H]^+$ and $[M+Na]^+$ and $[M+K]^+$ for the target peptide were the only species visible in the MALDI-TOF spectra of the crude product (**Figure 5.21**).



29-35
1063.9 Da

Figure 5.20: Chemical structure of the Dbz containing CCL2 peptide fragment Section **29-35.**

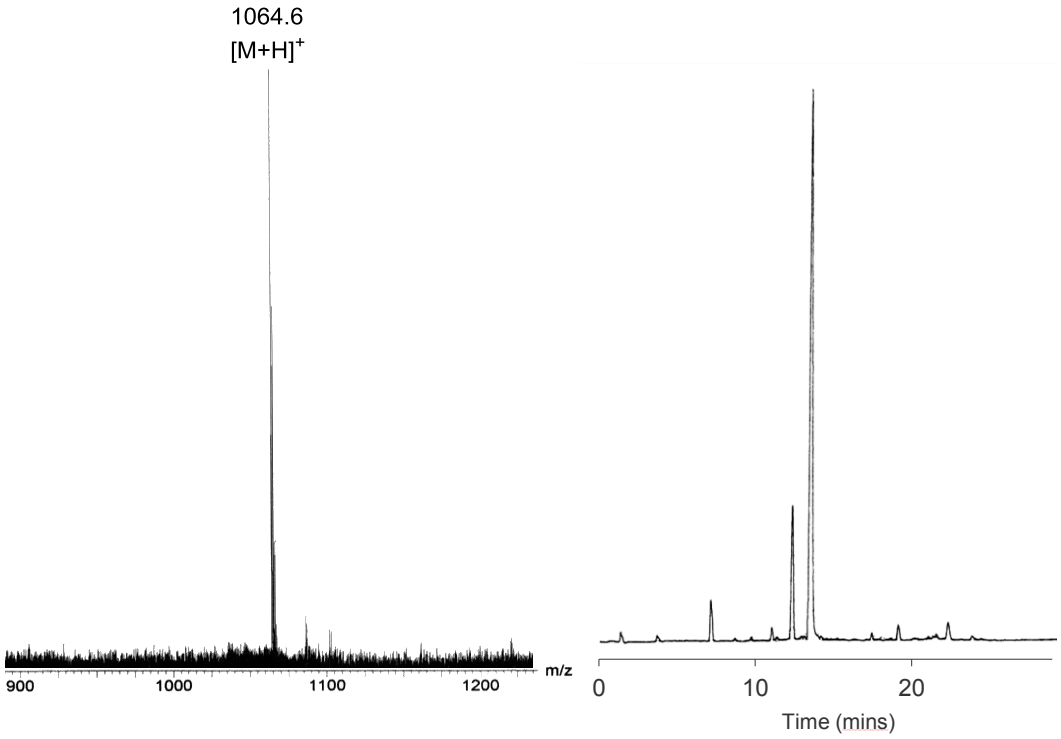
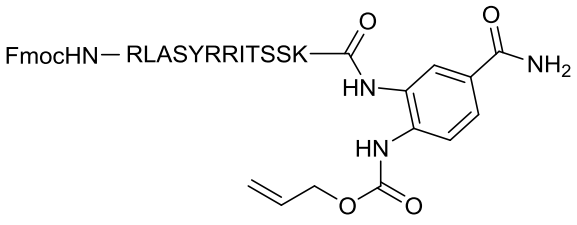


Figure 5.21: Analysis of the crude CCL2 peptide fragment **29-35** *Left*: positive ion MALDI-TOF spectrum. *Right*: Analytical RP-HPLC (230 nm). The linear gradient ranges from 2% MeCN (0.1% TFA) in H_2O (0.1% TFA) at 0 minutes to 50% MeCN (0.1% TFA) in H_2O (0.1% TFA) at 30 minutes.

The synthesis was continued to Arg24 (**Figure 5.22**) and RP-HPLC analysis showed that the crude peptide had a good level of purity (**Figure 5.23**). In addition, the $[M+H]^+$ for the target peptide was the only species visible in the MALDI-TOF spectra of the crude product (**Figure 5.23**).



24-35
1876.0 Da

Figure 5.22: Chemical structure of the Dbz containing CCL2 peptide fragment Section **24-35**. It should be noted that the *N*-Fmoc group was not removed prior to this test cleavage.

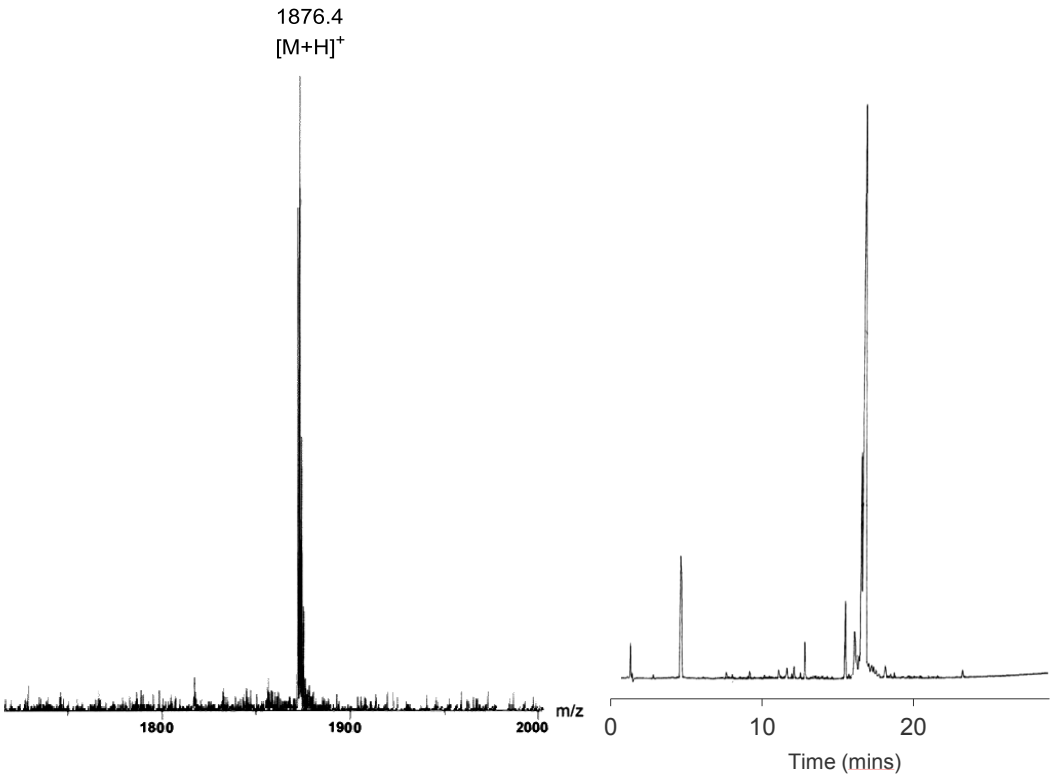
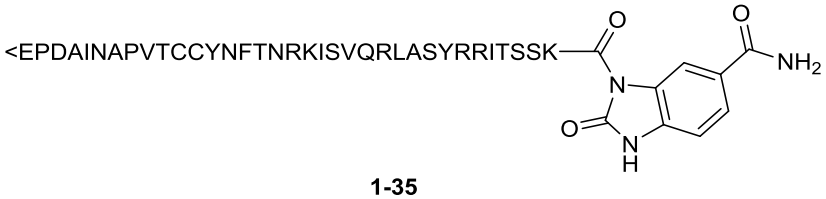


Figure 5.23: Analysis of the crude CCL2 peptide fragment **24-35**. *Left*: positive ion MALDI-TOF spectrum. *Right*: Analytical RP-HPLC (230 nm). The linear gradient ranges from 2% MeCN (0.1% TFA) in H_2O (0.1% TFA) at 0 minutes to 50% MeCN (0.1% TFA) in H_2O (0.1% TFA) at 30 minutes.

A sharp major HPLC peak and MALDI-TOF signal in their respective spectra are observed. Thus, the remaining procedure (24 AA couplings) was carried with no further test cleavages to complete the synthesis of the full CCL2 peptide fragment **1-35**. The post synthesis *N*-Alloc deprotection and Dbz cyclisation that lead to the final Nbz containing CCL2 peptide fragment **1-35** (**Figure 5.24**) were conducted manually at room temperature (**Table 5.8**).



1-35
4142.1 Da

Figure 5.24: Chemical structure of the Nbz containing CCL2 peptide fragment Section **1-35**.

The complete the synthesis of the full CCL2 peptide fragment **1-35** resulted in one major peak in the RP-HPLC spectrum (retention time ≈ 19 min, **Figure 5.25**) and LCMS-ESI$^+$ spectrum indicated that the peak was the correct mass with visible $[M+3H]^{3+}$, $[M+4H]^{4+}$and $[M+6H]^{6+}$ peaks. The crude peptide purity was 20% with a variety of early- and late-running impurities visible in the RP-HPLC spectrum and a number of unassignable impurities in the LCMS-ESI$^+$ spectrum. As in **Section 5.3.1.2**, peaks indicative of the hydrolysis of the Nbz are visible in the LCMS Spectrum (-Nbz).

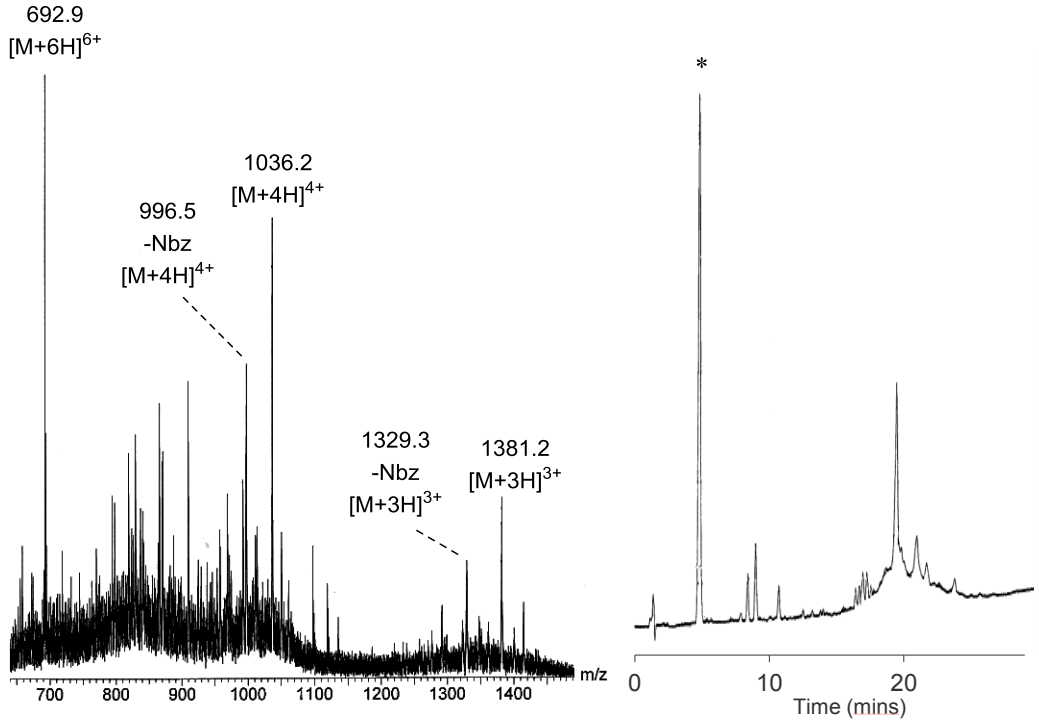


Figure 5.25: Analysis of the crude CCL2 peptide fragment **1-35**. *Left*: ESI+ spectrum. *Right*: Analytical RP-HPLC (230 nm). The linear gradient ranges from 2% MeCN (0.1% TFA) in H_2O (0.1% TFA) at 0 minutes to 50% MeCN (0.1% TFA) in H_2O (0.1% TFA) at 30 minutes.*indicates front-running small molecule impurity from Alloc deprotection.

Purification of CCL2 peptide fragment **1-35** proved exceptionally difficult and the purest fractions yielded product peptide at 67% and 75% purity (*left* and *right*, respectively in **Figure 5.26**) in 1% yield.

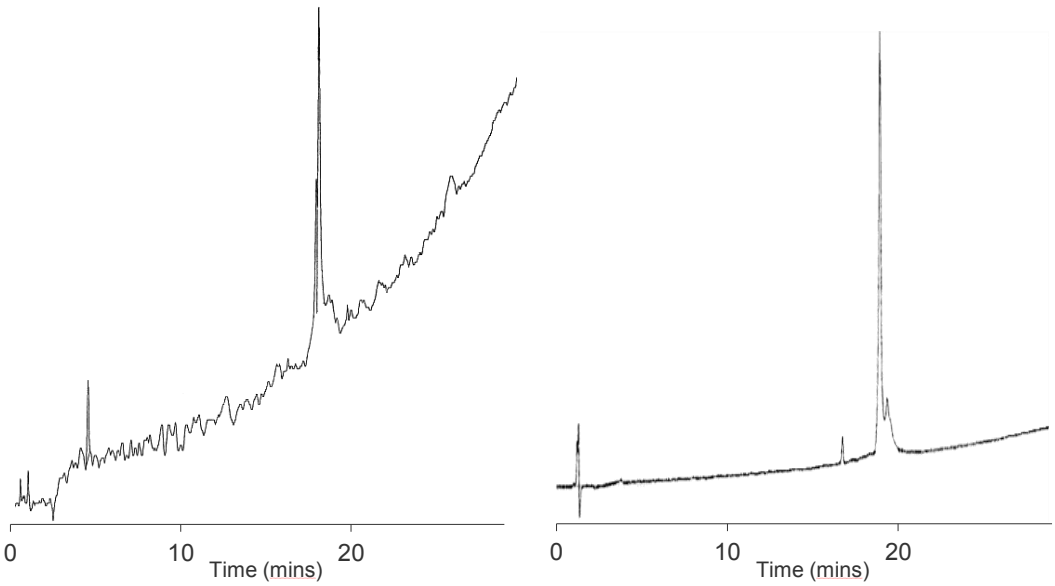


Figure 5.26: Analysis of the purified fragment **1-35**. *Left*: Analytical RP-HPLC (230 nm) of purified fragment 1-35, integration shows 67% pure product. The linear gradient ranges from 2% MeCN (0.1% TFA) in H_2O (0.1% TFA) at 0 minutes to 50% MeCN (0.1% TFA) in H_2O (0.1% TFA) at 30 minutes. *Right:* Analytical RP-HPLC (230 nm) of purified fragment **1-35**, integration shows 75% pure product. The linear gradient ranges from 2% MeCN (0.1% TFA) in H_2O (0.1% TFA) at 0 minutes to 50% MeCN (0.1% TFA) in H_2O (0.1% TFA) at 30 minutes.

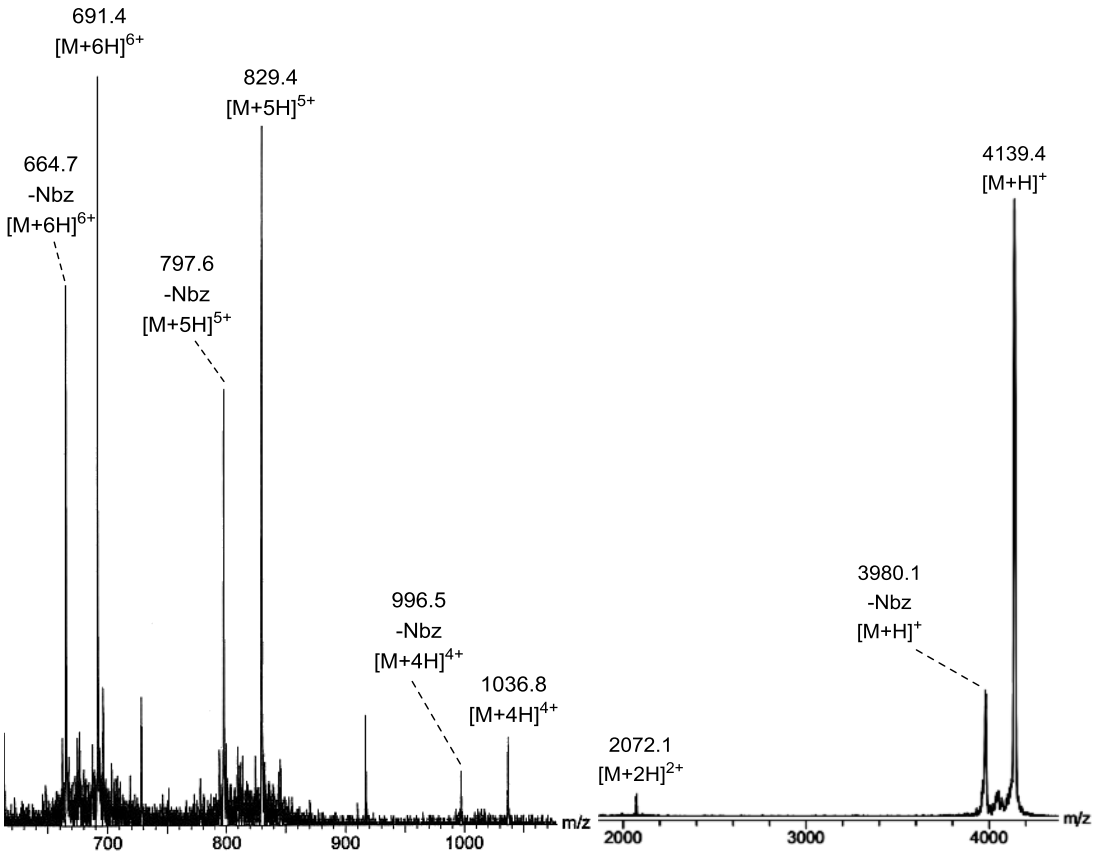
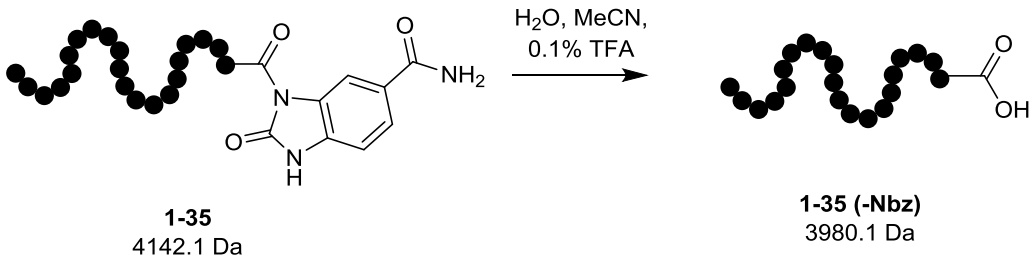


Figure 5.27: Analysis of the purified* CCL2 fragment **1-35**. *Left*: ESI+ spectrum. *Right*: positive ion MALDI-TOF spectrum. *Spectra correspond to 75% pure fraction (by analytical RP-HPLC (*right*, **Figure 5.26**))

The low purity was attributed to a significant amount of hydrolysis in a large number of fractions after purification. It was questioned whether the stability of the peptide fragment **1-35** (with C-terminal Lys-Nbz) in purification/analysis solvents (H_2O/MeCN with 0.1% TFA) was low and that degradation throughout the purification and analysis led to this (**Scheme 5.6**). To investigate this, an analytical RP-HPLC and MS were run after the pure product peptide was stored in solution at RT for 4 h (0.1% TFA and 20% MeCN in H_2O at RT). The RP-HPLC after solution storage (*right*, **Figure 5.28**) showed considerable peak broadening and the LCMS-ESI^+ mass spectrum (**Figure 5.29**) showed extensive degradation of the product peak. A number of unassignable peaks were observed as well as the $[M+3H]^{3+}$ and $[M+4H]^{4+}$ peaks corresponding to the hydrolysis (-Nbz) product at 1329.6 Da and 997.4 Da, respectively.



Scheme 5.6: Hydrolysis of CCL2 fragment **1-35.**

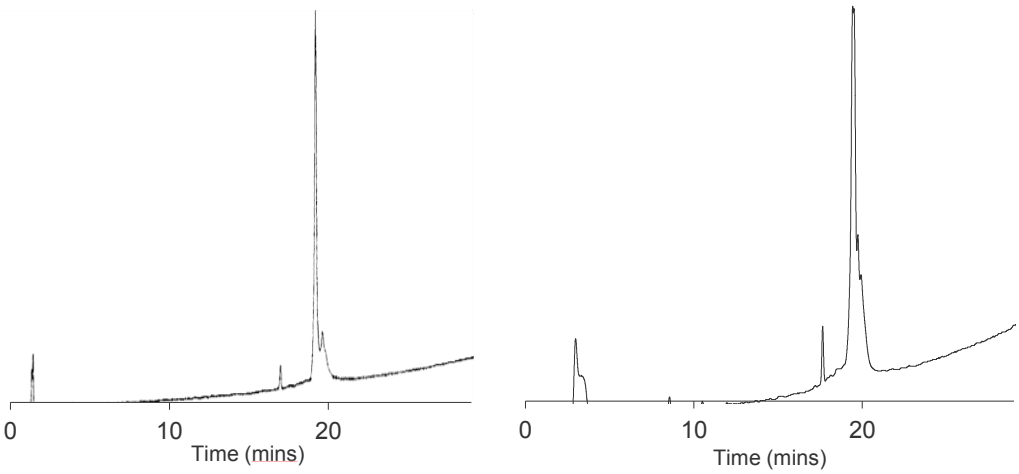


Figure 5.28: *Left* Analytical RP-HPLC (230 nm) of purified fragment 1-35. *Right*: Analytical RP-HPLC (230 nm) of purified fragment 1-35 after solution storage for 4 h. The linear gradients range from 2% MeCN (0.1% TFA) in H_2O (0.1% TFA) at 0 minutes to 50% MeCN (0.1% TFA) in H_2O (0.1% TFA) at 30 minutes.

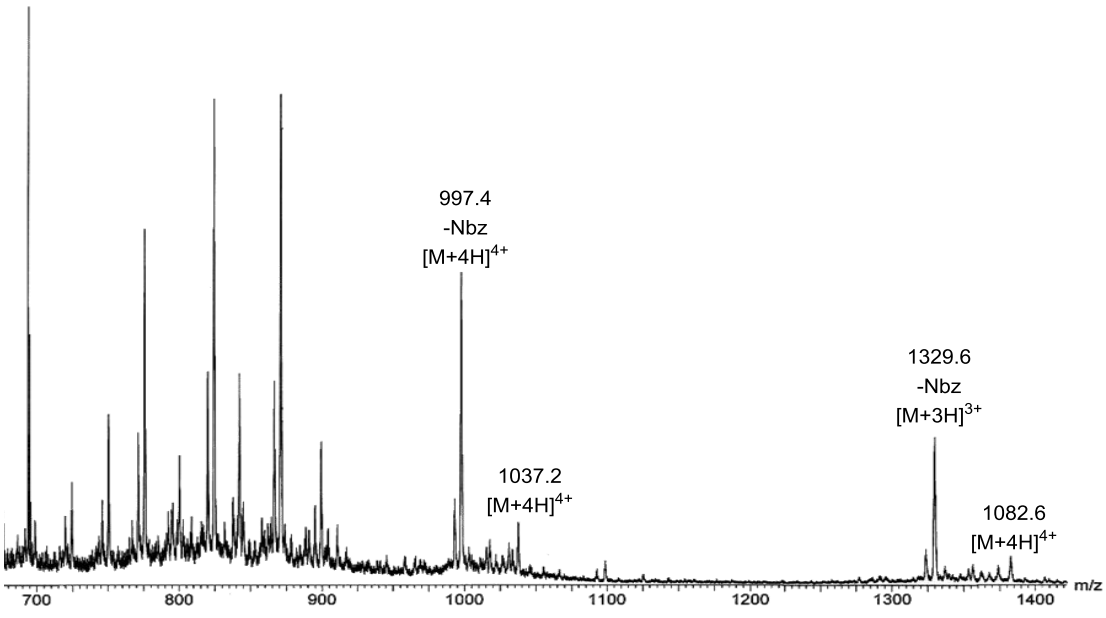


Figure 5.29: LCMS ESI+ spectrum of purified CCL2 peptide fragment **1-35** after solution storage for 4 h.

Due to the low solution stability of this peptide, the purified product was stored as a freeze-dried powder at -20 °C and only dissolved in solution when the NCL was undertaken.

5.4.3 NCL of CCL2 Peptide Fragment 1-35 and Fragment 36-76

With peptide fragments **1-35** and **36-76** in hand, an NCL reaction was carried out based on the Novabiochem procedure[14] on a small scale (1.5 mg) (*i* in **Figure 5.30**).

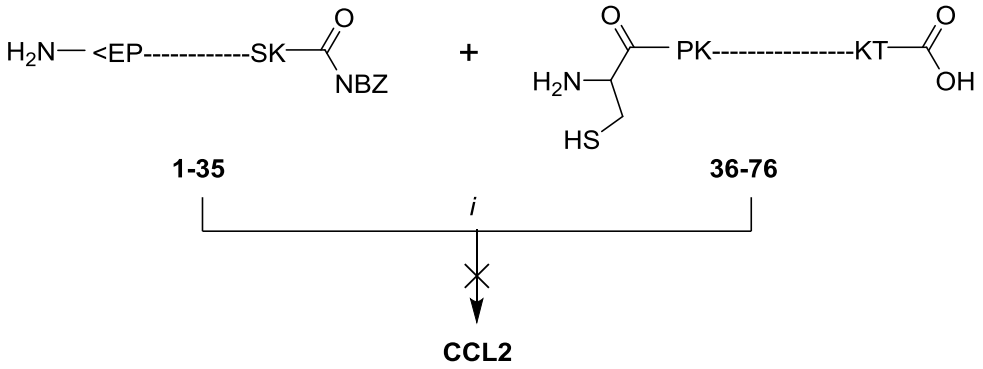


Figure 5.30: Attempted NCL of CCL2 peptide fragments **1-35** and **36-76** to form CCL2. *i:* MPAA, sodium phosphate, TCEP, Gdn HCl, H_2O, RT, 24 h.

No trace of the full length ligation product (CCL2) was observed at 4, 8, or 24 h in MALDI-TOF or LCMS spectra (24 h, MALDI-TOF, **Figure 5.31**). After 24 h the reaction was purified *via* size exclusion chromatography and RP-HPLC. No discernible CCL2 product peaks were observed in the MALDI-TOF and/or LCMS-ESI$^+$ spectra of any fraction.

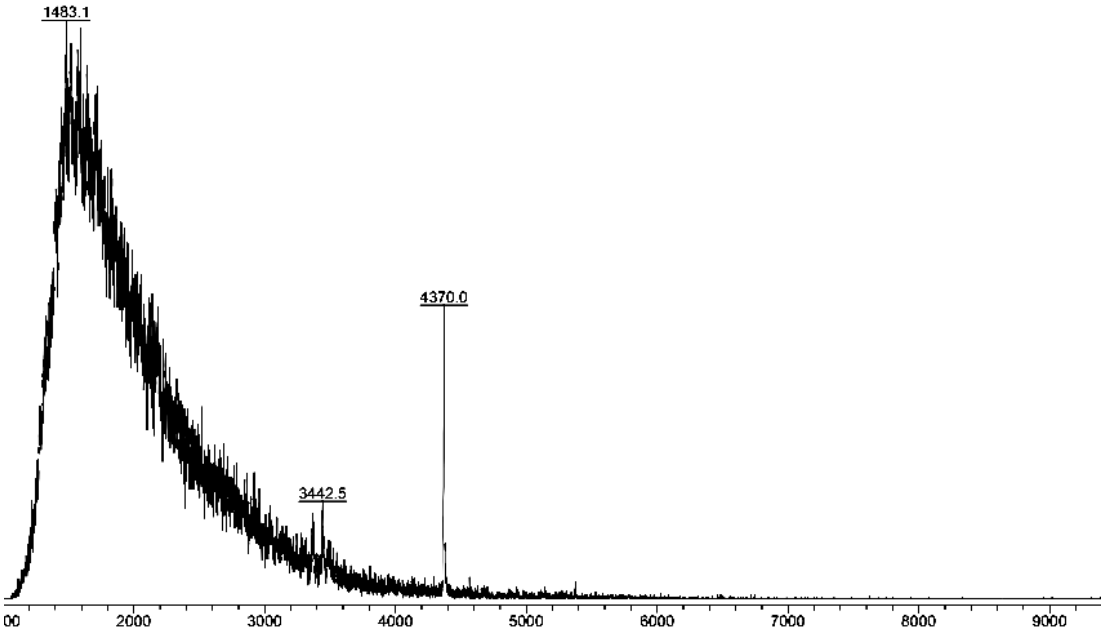


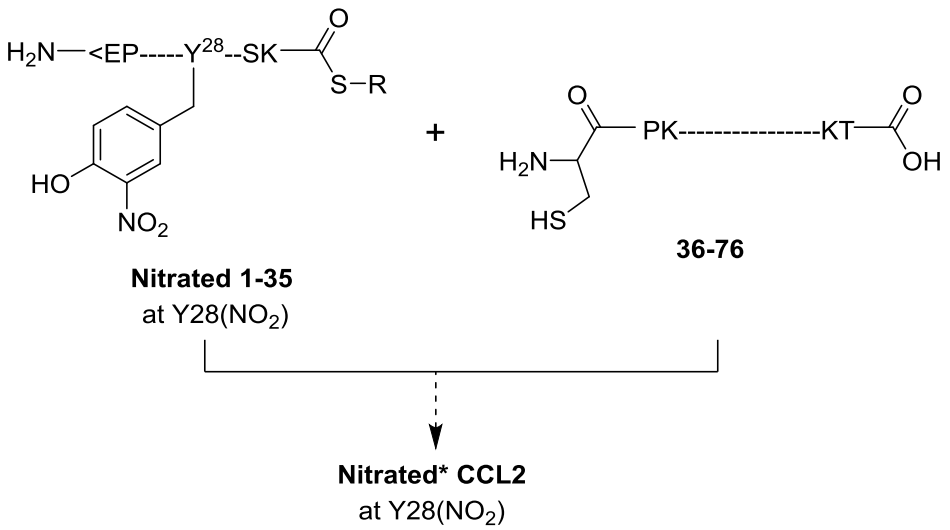
Figure 5.31: MALDI-TOF spectra of the crude NCL of CCL2 peptide fragments **1-35** and **36-76** after 24 h.

5.5 Synthesis of Nitrated Tyrosine CCL2

5.5.1 Incorporation of a Nitrated Tyrosine in CCL2

Although the nitration of CCL2 is possible and nitrated variants have been produced and these derivatives are known to abrogate that ability of CCL2 to initiate chemotaxis (**Chapter 1**). In all of the derivatives synthesized the nitration modification is not site-specific. Hence, the abrogating effect of nitration on activity cannot be pinpointed to a particular residue or mechanism of inhibition. Therefore it is our aim to adapt the strategy that was attempted for the synthesis native CCL2 to synthesize a small library of site-specifically nitrated CCL2 derivatives. For

example, the Tyr(NO_2)28 nitrated CCl2 derivative would be synthesised following the procedure to synthesize CCL2 peptide fragment **1-35** with the Fmoc-L-3-nitrotyrosine-OH building block replacing the standard Fmoc-Tyr(O*t*Bu)-OH at the relevant position. The use of Fmoc-L-3-nitrotyrosine-OH in microwave assisted SPPS is previously described.[42] With a nitrated CCL2 fragment **1-35** in hand, a subsequent NCL with CCL2 peptide fragment **36-76** (already prepared in **Section 5.4.1**) to yield the desired product could be envisaged (**Scheme 5.7**).

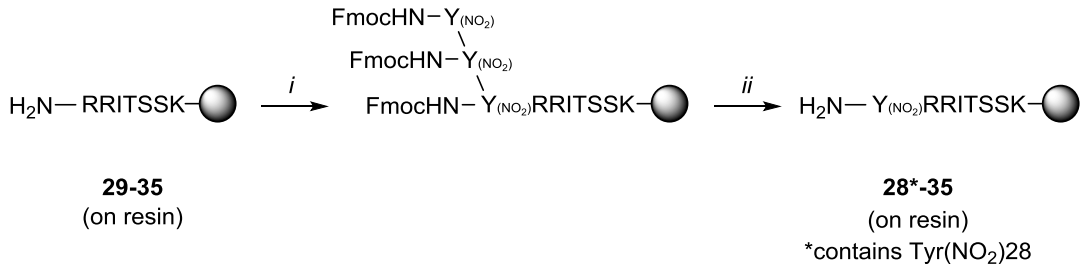


Scheme 5.7: Proposed synthesis of site-specifically (Y28(NO_2)) nitrated CCL2 *via* an NCL of CCL2 peptide fragment **36-76** and nitrated (Y28(NO_2)) fragment 1-35.

Whilst the synthesis of CCL2 peptide fragment **1-35** was underway a small section of the resin that was synthesised up to Arg29 (**Figure 5.20** and **Figure 5.21**, **Section 5.4.2**) was used in an attempt to synthesize nitrated versions of fragment **1-35** containing Tyr28(NO_2). In part, this work was carried out to prove the Fmoc-L-3-nitrotyrosine-OH building block was stable under our microwave assisted SPPS techniques.

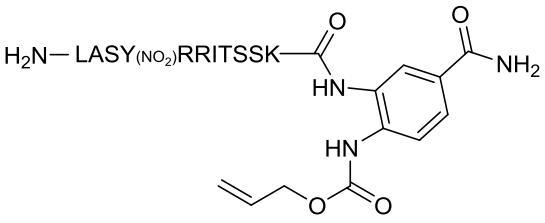
The 3-nitrotyrosine residue was coupled (**Scheme 5.8**) to Alloc protected Dbz peptide fragment **29-35** (pre-made in **Section 5.4.2**) using manual room temperature coupling conditions. These conditions were chosen so that colorimetric tests (TNBS, Kaiser and chloranil) could indicate the reaction efficiency. However, this was not possible as the coupling of Tyr(NO_2) to the resin reaction caused severe and permanent discoloration of the resin beads, rendering these tests impossible.

Moreover, significant quantities of mono- di and tri-addition products were observed during the peptide coupling (*i* in **Scheme 5.8**) yet treatment with the deprotection conditions (*ii* in **Scheme 5.8**) yielded the desired production: nitrated (Tyr28) CCL2 peptide fragment **28-35**. The multi-addition products arise because the Fmoc-L-3-nitrotyrosine-OH contains no protection on the phenoxy group, hence, the phenolic OH can attack activated acids (which are in excess) and form an ester. The ester formed is unstable to the basic conditions used in the Fmoc deprotection step and is cleaved.[37]



Scheme 5.8: Manual coupling of Fmoc-L-3-nitrotyrosine(OH)-OH onto CCL2 peptide fragment **29-35** on resin (**Section 5.4.2**). *i*: Fmoc-L-3-nitrotyrosine(OH)-OH (5 eq), PyBOP (5 eq), DIPEA (10 eq), DMF, 1 h 30 min, RT, x2. *ii*: piperidine, DMF, RT, 5 min; piperidine, DMF, RT, 15 min.

The synthesis was continued for a further 3 couplings up to Leu25 (**Figure 5.32**) and RP-HPLC analysis showed that the crude peptide had a good level of purity (**Figure 5.33**). In addition, characteristic[43] major peaks were observed in the MALDI-TOF (**Figure 5.33)** indicating that the synthesis of the nitro-peptide was effective. However, a serine deletion (again with characteristic nitro-tyrosine peak) was observed at 1456.8 Da (**Figure 5.33**). The serine deletion was not observed previously and thus, must arise due to incomplete coupling of serine to the N-terminal 3-nitrotyrosine. Tentatively, we can deduce that the coupling onto a N-terminal 3-nitrotyrosine can be deemed a difficult coupling and future syntheses would take this into account. For example, coupling at higher temperature or triple coupling as well as the preparation of a (currently non-commercially available) hydroxyl protected Tyr(NO_2) could overcome this problem.



25-35*
1544.4 Da
*contains $Tyr(NO_2)$28

Figure 5.32: Chemical structure of the Dbz containing nitrated CCL2 peptide fragment **25-35**.

.

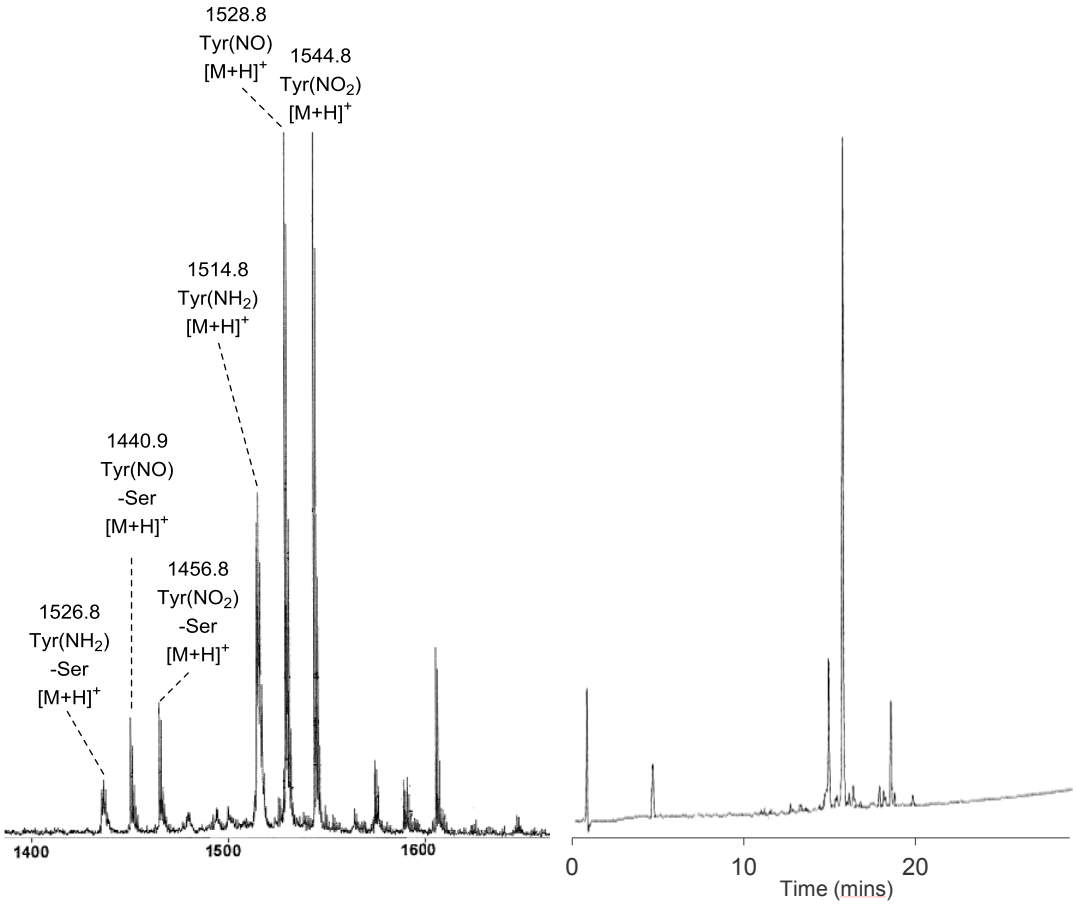


Figure 5.33: Analysis of the crude Dbz containing nitrated CCL2 peptide fragment **25-35**. *Left*: MALDI-TOF spectrum. *Right*: Analytical RP-HPLC (230 nm). The linear gradient ranges from 2% MeCN (0.1% TFA) in H_2O (0.1% TFA) at 0 minutes to 50% MeCN (0.1% TFA) in H_2O (0.1% TFA) at 30 minutes.

5.6 Conclusions

This work aimed to develop the first microwave assisted SPPS of native CCL2 to enable the fast production of CCL2 and a number of site-selectively nitrated analogues. Previously in the Cobb group the full linear microwave assisted synthesis of CCL2 was attempted and only a CCL2 peptide fragment of 25 amino acids in length could be produced in purity. A number of problems were overcome (e.g. progressing through regions of aggregation and methionine oxidation) and an optimised synthesis was developed that pushed the synthesis of CCL2 up to a peptide fragment (**36-76**) of 41 amino acids in length. The synthesis of CCL2 peptide fragment **36-76** was very successful and over tens of mgs of pure peptide were produced in excellent purity from hundreds of mgs of crude, with only 20% of the material synthesised being used. However, the full microwave linear SPPS of CCL2 was still not possible. Therefore, CCL2 peptide fragment (**1-35**) was synthesised and an NCL reaction attempted to form the full sequence.

The syntheses of CCL2 peptide fragment **1-35** was successful and test cleavages throughout showed good purity by analytical HPLC and mass spectra. However, the post-synthesis purification, analysis, isolation of the CCL2 peptide fragment **1-35** (N-terminal Nbz peptide) was achieved but only a small amount of partially pure material was produced. The subsequent NCL with CCL2 peptide fragment **36-76** was attempted but analysis did not show any of the desired product. Further optimisation of the NCL reaction was not feasible without re-synthesis as all the partially pure material produced was used. However, to extend this work and produce CCL2 (*via* fragmented microwave assisted SPPS) the use of alternative resin to Dawson Dbz would be envisaged (e.g. Sulfamylbutyryl or SEA, **Table 5.1**).

Trial syntheses of a modified analogue confirmed the Fmoc-L-3-nitrotyrosine-OH building block to be amendable to microwave assisted SPPS, in our hands. Thus, if an improved route to the synthesis of a CCL2 peptide fragment **1-35** (containing a thioester or equivalent) was afforded then the synthesis of tyrosine nitrated derivatives could swiftly follow as the coupling partner (CCL2 peptide fragment **36-76**) is already in hand.

5.7 References

1. K. Kim and P. A. Cole, *J. Am. Chem. Soc.*, 1998, **120**, 6851-6858.
2. J. Pepin, C. Guern, F. Milord and P. J. Schechter, *Lancet*, 1987, **330**, 1431-1433.
3. C. Frieden, S. D. Hoeltzli and J. G. Bann, *Methods Enzymol.*, 2004, **380**, 400-415.
4. N. K. Mishra, A. K. Urick, S. W. J. Ember, E. Schönbrunn and W. C. Pomerantz, *ACS Chem. Biol.*, 2014, **9**, 2755-2760.
5. C. Murphy, B. Clark and J. Amadio, *Appl. Microbiol. Biotechnol.*, 2009, **84**, 617-629.
6. M. H. Gelb, J. P. Svaren and R. H. Abeles, *Biochemistry*, 1985, **24**, 1813-1817.
7. S. K. Holmgren, K. M. Taylor, L. E. Bretscher and R. T. Raines, *Nature*, 1998, **392**, 666-667.
8. T. Allmendinger, E. Felder and E. Hungarbühler, *Tetrahedron Lett.*, 1990, **31**, 7301-7304.
9. J. B. Tuttle, J. M. Azzarelli, B. M. Bechle, A. B. Dounay, E. Evrard, X. Gan, S. Ghosh, J. Henderson, J.-Y. Kim, V. D. Parikh and P. R. Verhoest, *Tetrahedron Lett.*, 2011, **52**, 5211-5213.
10. N. K. O'Connor, D. K. Rai, B. R. Clark and C. D. Murphy, *J. Fluor. Chem.*, 2012, **143**, 210-215.
11. P. Van der Veken, K. Senten, I. Kertèsz, I. De Meester, A.-M. Lambeir, M.-B. Maes, S. Scharpé, A. Haemers and K. Augustyns, *J. Med. Chem.*, 2004, **48**, 1768-1780.
12. S. Moran, D. K. Rai, B. R. Clark and C. D. Murphy, *Org. Biomol. Chem.*, 2009, **7**, 644-646.
13. M. D. Liptak, K. C. Gross, P. G. Seybold, S. Feldgus and G. C. Shields, *J. Am. Chem. Soc.*, 2002, **124**, 6421-6427.
14. L. Volpon, F. Besson and J.-M. Lancelin, *Eur. J. Biochem.*, 1999, **264**, 200-210.
15. U. Schmidt, A. Lieberknecht and J. Wild, *Synthesis*, 1984, 53-60.
16. N. K. O'Connor, A. S. Hudson, S. L. Cobb, D. O'Neil, J. Robertson, V. Duncan and C. D. Murphy, *Amino Acids*, 2014, **46**, 2745-2752.
17. A. S. K. Hashmi, P. Haufe, C. Schmid, A. Rivas Nass and W. Frey, *Chem. - Eur. J.*, 2006, **12**, 5376-5382.
18. M.-H. Seo, J. Han, Z. Jin, D.-W. Lee, H.-S. Park and H.-S. Kim, *Anal. Chem.*, 2011, **83**, 2841-2845.
19. L. Isakovic, O. M. Saavedra, D. B. Llewellyn, S. Claridge, L. Zhan, N. Bernstein, A. Vaisburg, N. Elowe, A. J. Petschner, J. Rahil, N. Beaulieu, F. Gauthier, A. R. MacLeod, D. Delorme, J. M. Besterman and A. Wahhab, *Bioorg. Med. Chem. Lett.*, 2009, **19**, 2742-2746.
20. E. Negishi, *Acc. Chem. Res.*, 1982, **15**, 340-348.
21. Y. Berger, H. Dehmlow, D. Blum-Kaelin, E. A. Kitas, B.-M. Löffler, J. D. Aebi and L. Juillerat-Jeanneret, *J. Med. Chem.*, 2004, **48**, 483-498.
22. L. P. Partida-Martinez, C. Flores de Looß, K. Ishida, M. Ishida, M. Roth, K. Buder and C. Hertweck, *Appl. Envirom. Microb.*, 2007, **73**, 793-797.
23. I. Rilatt, L. Caggiano and R. F. W. Jackson, *Synlett*, 2005, 2701-2719.
24. C. L. Oswald, T. Carrillo-Márquez, L. Caggiano and R. F. W. Jackson, *Tetrahedron*, 2008, **64**, 681-687.

25. M. F. Eerland and C. Hedberg, *J. Org. Chem.*, 2012, **77**, 2047-2052.
26. Y. Belokon, V. Maleev, T. Savel'eva, M. Moskalenko, D. Pripadchev, V. Khrustalev and A. Saghiyan, *Amino Acids*, 2010, **39**, 1171-1176.
27. H. J. C. Deboves, C. A. G. N. Montalbetti and R. F. W. Jackson, *J. Chem. Soc., Perkin Trans. 1*, 2001, 1876-1884.
28. R. F. W. Jackson, I. Rilatt and P. J. Murray, *Org. Biomol. Chem.*, 2004, **2**, 110-113.
29. S. L. Cobb and J. C. Vederas, *Org Biomol Chem*, 2007, **5**, 1031-1038.
30. H. M. Wisniewska, E. C. Swift and E. R. Jarvo, *J. Am. Chem. Soc.*, 2013, **135**, 9083-9090.
31. V. B. Phapale and D. J. Cardenas, *Chem. Soc. Rev.*, 2009, **38**, 1598-1607.
32. G. Manolikakes, C. Muñoz Hernandez, M. A. Schade, A. Metzger and P. Knochel, *J. Org. Chem.*, 2008, **73**, 8422-8436.
33. R. F. W. Jackson, M. J. Wythes and A. Wood, *Tetrahedron Lett.*, 1989, **30**, 5941-5944.
34. R. Caputo, M. DellaGreca, I. de Paola, D. Mastroianni and L. Longobardo, *Amino Acids*, 2010, **38**, 305-310.
35. V. Eswarakrishnan and L. Field, *J. Org. Chem.*, 1981, **46**, 4182-4187.
36. R. Filler and R. Saha, *Future Med. Chem.*, 2009, **1**, 777-791.
37. C. L. Weeks, A. Polishchuk, Z. Getahun, W. F. Degrado and T. G. Spiro, *J. Raman Spectrosc.*, 2008, **39**, 1606-1613.
38. P. C. Cirino, Y. Tang, K. Takahashi, D. A. Tirrell and F. H. Arnold, *Biotechnol. Bioeng.*, 2003, **83**, 729-734.
39. M. D. Crespo and M. Rubini, *PLOS ONE*, 2011, **6**, e19425.
40. L. Merkel, H. S. G. Beckmann, V. Wittmann and N. Budisa, *Chembiochem*, 2008, **9**, 1220-1224.
41. J. Wang, M. Sánchez-Roselló, J. L. Aceña, C. del Pozo, A. E. Sorochinsky, S. Fustero, V. A. Soloshonok and H. Liu, *Chem. Rev.*, 2013, **114**, 2432-2506.
42. M. Monclus, C. Masson and A. Luxen, *J. Fluor. Chem.*, 1995, **70**, 39-43.
43. R. L. VonTersch, F. Secundo, R. S. Phillips and M. G. Newton, *Biomedical Frontiers of Fluorine Chemistry*, American Chemical Society, 1996.

Chapter 6 : Conclusions and Future Work

6.1 Conclusions and Future Work

This work describes the syntheses of a variety of amino acids (**Section 6.1.1**), cyclic dipeptides/DKPs (**Section 6.1.2**) and large linear CCL2 fragment peptides (**Section 6.1.3**). Of the selection of amino acids and dipeptides synthesised, a number have been utilised in chemical and biological applications. The linear peptides are intermediates in the total microwave assisted SPPS of CCL2, and, although this was not completed, progress was made towards developing a viable synthetic route.

6.1.1 Amino Acid Syntheses

A library of heteroaromatic amino acids has been prepared in moderate yields *via* an optimised Negishi cross-coupling reaction (**Figure 6.1**). Previously, the synthesis of these heteroaromatic amino acids was not described *via* this route with only one *N*-Boc protected derivative (of **16**) reported in the literature. The developed route offers an effective alternate strategy to access heteroaromatic amino acids from a simple starting material (iodoalanine). The yields are comparable to other more established routes (e.g. asymmetric dehydrogenation, chiral glycine equivalents and enzymatic synthesis) and in some cases may be preferable than other strategies. For example, in asymmetric hydrogenation the *ee* of desired product can be extremely dependant on having *N*- and *C*- protecting groups with steric bulk.

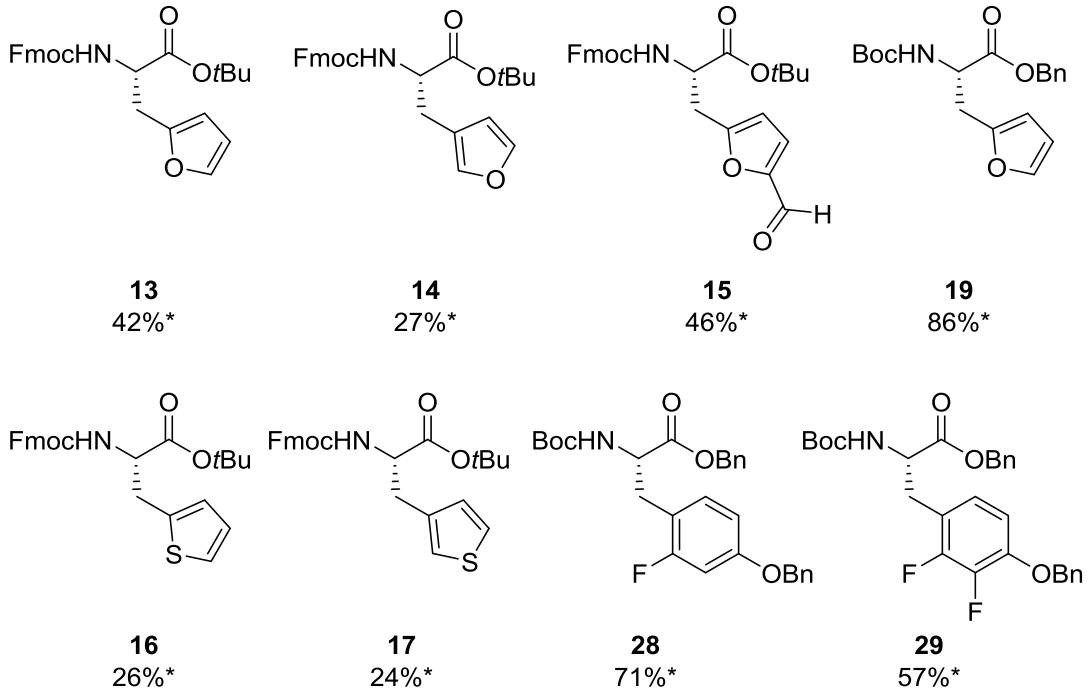
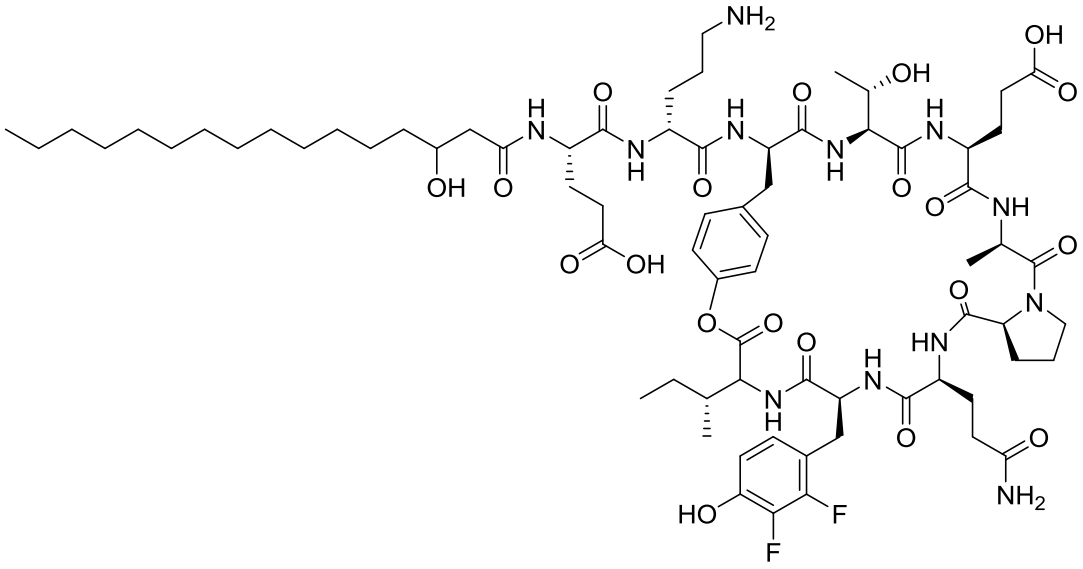


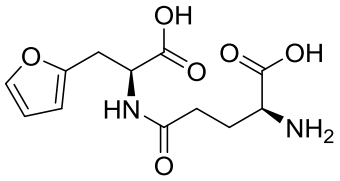
Figure 6.1: The various heteroaromatic amino acids reported in this work synthesized via Pd-catalysed Negishi cross-coupling (**Chapter 2**). *denotes highest isolated yield of the amino acid.

Some of the heteroaromatic amino acid building blocks were found to be tolerant to incorporation into bioactive molecules *via* chemical and biological syntheses. For example fluorinated 2-3-difluorotyrosine (**29**) was shown to be amenable to biosynthetic incorporation into the lipopeptide fengycin (**Figure 6.2**). Furylalanine was shown to be stable to solution and SPPS, thus, can be easily incorporated into bioactive peptides e.g. the attempted syntheses of γ-glutamyl-2-furylalanine (**Figure 6.3**).



Fluorinated Fengycin

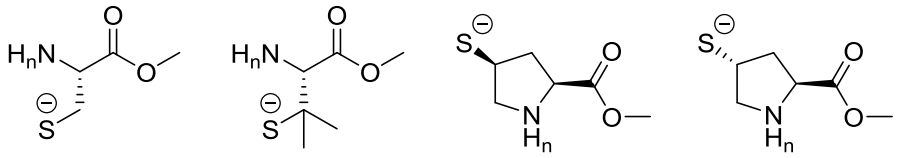
Figure 6.2: A fengycin derivative lipopeptide containing a 2,3-difluorotyrosine derivative amino acid.



y-glutamyl-2-furylalanine.

Figure 6.3: The chemical structure of natural product: γ-glutamyl-2-furylalanine.

A small library of thiol containing amino acids (**31**, **32**, **33** and **34**) was produced. This work included the synthesis of two different stereo-isomers of 4-substituted thio-proline *via* multiple synthetic routes. The thiol containing amino acids (**33** and **34**) were then used to provide information (in the form of pK_a values and ionic species abundance (**Figure 6.4**)) that could indicate their reactivity towards NCL and thus, be beneficial in the development of kinetic NCL reactions.



	31	**32**	**33**	**34**
pH 6.8	15%	8%	21%	-
pH 7.2	22%	13%	24%	-

Figure 6.4: The abundance of thiolate species (either zwitterionic or anionic) present at pH 6.8 and 7.2. The thiolate species is thought to be responsible for initial nucleophilic attack upon the thioester.***34** provided erroneous results due to oxidation.

The results show promise as a significant difference in the abundance of active deprotonated thiol species is observed over a small pH range (6.8 – 7.2). However, re-testing of all four amino acids under reducing conditions is needed so that data can be gathered on **34**. This result is of interest as a major difference in reactivity in NCL has been shown between **33** and **34**. To further this work, small peptides containing the terminal thiol amino acids could be synthesised and the utilised in competition experiments to correlate pK_a data and NCL efficiency. Ultimately, the correlated pK_a data could lead to facile planning and execution of kinetic "one-pot" ligations with these amino acids. If successful, the work could be easily extended to a larger library of thiol-amino acids.

The aforementioned amino acids could be envisaged to be used in a variety of other peptide based molecules including the synthesis of potential CCL2 induced chemotaxis inhibiting DKPs (**Section 6.1.2**) and the incorporation into large peptides or proteins (**Section 6.1.3**).

6.1.2 DKPs and CCL2 Induced Chemotaxis Inhibition

Previous work by the Cobb group had shown potential in the use of DKPs as selective inhibitors of CCL2 induced chemotaxis. Therefore, we sought to develop an improved synthetic route and utilise this to synthesise (and analyse) a number of DKPs so that inhibitors could be intelligently designed and their mechanism of inhibition probed.

A solid-phase route to the synthesis of the target DKPs was successfully optimised to enable the production of DKPs with a significant improvement on the previous solution phase approaches. Using this solid-phase approach a single DKP could be synthesised and purified within 8 hours, compared to the solution phase synthesis that could take several days, multiple purifications and high temperatures.

In total, eleven DKPs were synthesised (from *N*-Boc amino acid building blocks) and underwent biological testing against inhibition of CCL2 induced chemotaxis. Of the eleven DKPs only three exhibited inhibition of <40% inhibition at a concentration of 100 μM (red, **Figure 6.5**) and no inhibitors were found to be significantly better than those previously described (e.g. cyclo(L-Phe-L-Pro) black, **Figure 6.5**). Although it was not possible to obtain crystal structures of every DKP synthesized, preliminary analysis has shown that the less active molecules (e.g. cyclo(L-Tyr-L-Pro) and cyclo(L-Pro-D-Phe)) exhibit folding of the aromatic over the core 6-membered ring (common in aromatic DKPs of this type) and the more active molecules (e.g. cyclo(L-Phe-L-Pro) and cyclo(*p*-fluoro-L-Phe-L-Pro)) do not. Moreover, the work provided insights into the effect on the inhibition activity of changing stereochemistry, aromatic moieties and prolyl moieties. For example, potent inhibition of CCL2 induced chemotaxis is observed from both cyclo(L-Pro-L-Trp) and cyclo(L-Phe-L-Hyp) (red, **Figure 6.5**). This suggests that the scaffold is tolerant to variations in heteroaromatic group and proline substitution at the 4- position. Therefore, work is currently underway to produce a 3rd generation library of DKPs using a range of heteroaromatic and proline substituted amino acids.

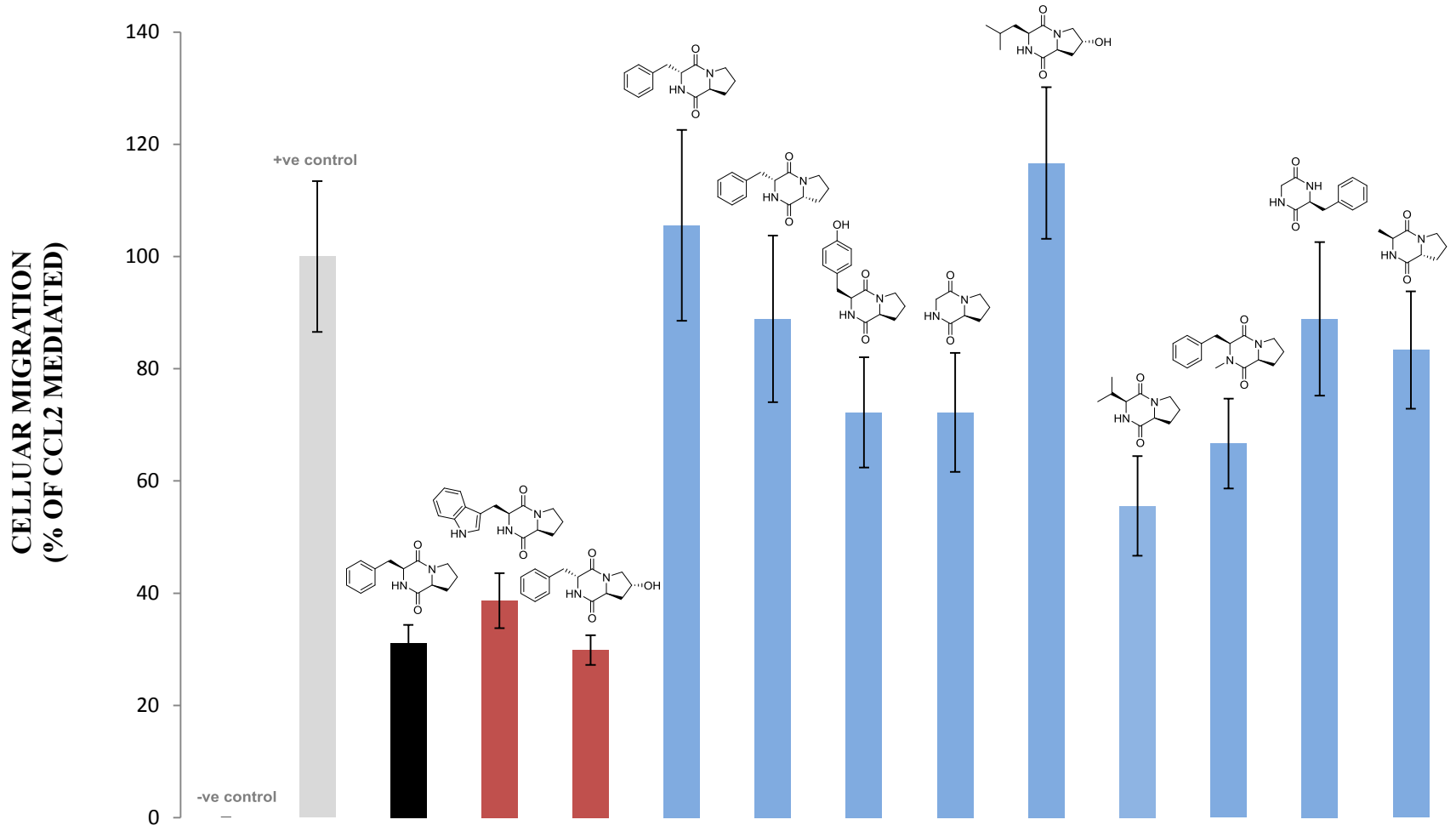
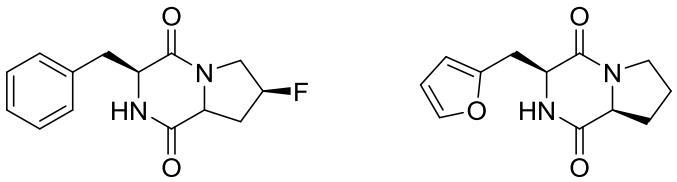


Figure 6.5: Testing of 100 µM DKPs (structures shown) against CCL2 mediated chemotaxis of THP-1 cells at 10 nM CCL2 concentration.

Two DKPs (**66** and **67, Figure 6.6**) containing non-natural amino acids were synthesised *via* the optimised solid-phase route. The previously described synthesis of the non-natural amino acid building blocks (4-fluoro proline, and 2-furyl alanine) facilitated this. The synthesis of a number of other non-natural amino acids has been described in this work and their application in the preparation of a new library of DKPs (containing **66** and **67**) is currently under investigation by the Cobb group (**Figure 6.7**).



66 **67**

Figure 6.6: Two DKPs (**66** and **67**) synthesised from the non-natural building block amino acids previously outlined.

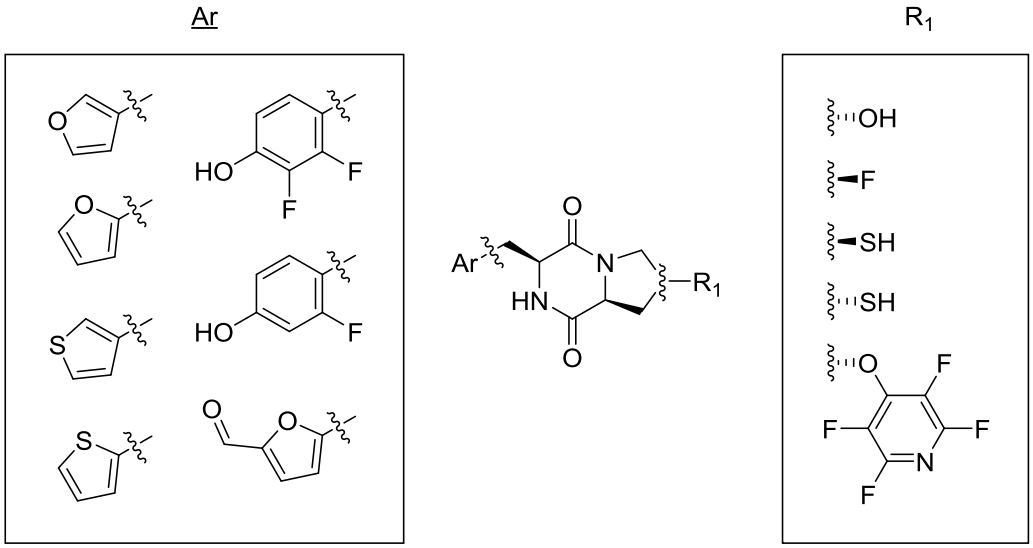


Figure 6.7: Generic structure of an aromatic/prolyl DKP (centre) and the possible substituent aromatic (Ar) and 4-substituted proline (R_1) groups that have had their synthesis reported in this work.

6.1.3 Potential Routes to the Microwave Assisted Synthesis of CCL2

Previous work by the Ali group at Newcastle University and others had shown that a heterogeneous mixture of nitrated (at Tyr13, Tyr28 and Trp59) chemokine CCL2 had reduced chemotactic activity and heparin binding; inferring that this posttranslational modification could be vital in abrogation of the immune response *in vivo*. To investigate this we set out to develop the first microwave assisted SPPS of native CCL2 (76 AA) to enable the fast production of CCL2 and a number of modified analogues. For example, CCL2 analogues containing nitrated or fluorinated tyrosine residues.

Previous work in the Cobb group had attempted the microwave assisted full linear SPPS, but the synthesis could not progress beyond the synthesis of a 25 AA CCL2 peptide fragment **52-76**. Significant improvements were made to this method and problems with aggregation and methionine oxidation were overcome to enable the synthesis of a 41 AA CCL2 peptide fragment **36-76**. However, the total linear SPPS remained out of reach. Ultimately, we were led to develop an SPPS of CCL2 that involved two fragments (**1-35** and **36-76**) and a NCL between them (**Figure 6.8**).

CCL2 Peptide fragment **1-35** was produced *via* microwave assisted SPPS on a Dawson Dbz resin. The resin was protected with *N*-Alloc (to prevent unwanted side-reactions formed under the high temperature microwave conditions) and the synthesis was shown to be effective throughout (*via* test cleavage followed by analytical HPLC/mass spectrometry). However, due to post-synthesis difficulties (in purification, analysis, isolation and subsequent reactions) of this fragment (containing an Nbz urea) the total synthesis of CCL2 was not completed. To extend and complete this work, it was envisaged that the CCL2 Peptide fragment **1-35** (and nitrated analogues, **Figure 6.8**) could be re-synthesised using a different resin and ligated with (already in hand) CCL2 Peptide fragment **36-76** to produce a library of site-selective CCL2 analogues.

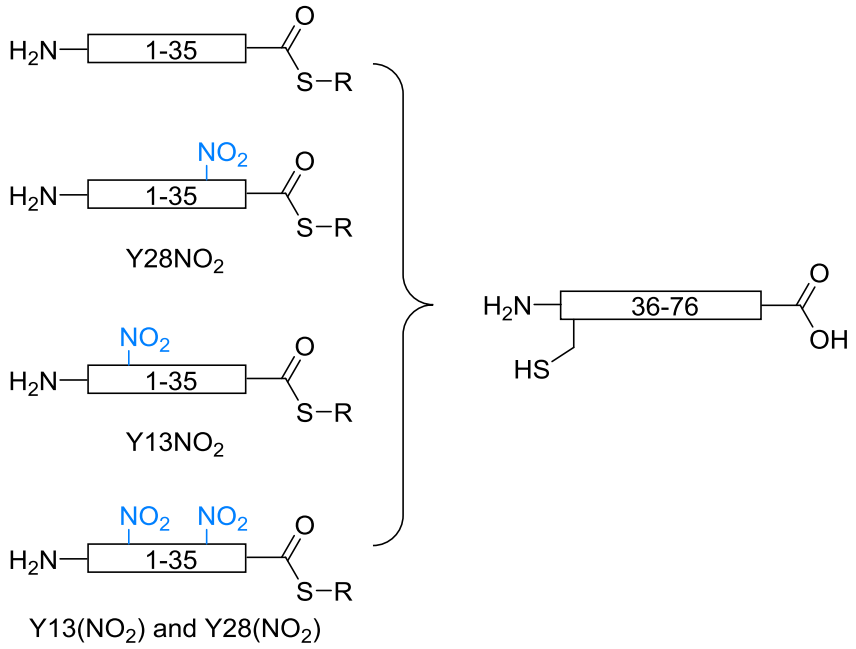


Figure 6.8: Proposed synthesis of site-specific nitrated variants of human chemokine: CCL2.

Chapter 7 : Experimental

7.1 Small Molecule Syntheses

7.1.1 Materials and Methods

All reported yields refer to the isolated yield and the product purity was estimated to be >95% by ^{1}H NMR. IR spectra were recorded on a Perkin Elmer Spectrum RX1 fitted with an ATR attachment. ^{1}H NMR spectra were recorded at 400 MHz using a Bruker Avance 400 MHz and ^{13}C NMR spectra at 100 MHz with a Bruker Avance. Chemical shifts are reported in ppm and are referenced to residual solvent peaks; $CHCl_3$ (^{1}H 7.26 ppm, ^{13}C 77.0 ppm). *J* couplings are measured in Hertz (Hz). 2D NMR techniques (COSY and HSQC) were used when chemical assignment of NMR spectra is given. Mass spectra were collected on a Waters TQD mass spectrometer and accurate mass spectra were collected on a Waters LCT Premier XE mass spectrometer. High pressure liquid chromatography was performed on an analytical Varian LC with a diode array detector. Optical rotations were measured with a Jasco P-1020 polarimeter. All reactions were monitored by T.L.C. using Merck pre-coated silica gel plates, Column chromatography was performed using silica gel (40-60 mM) and the specific solvent system indicated in the experimental procedures. All chemicals were purchased from Sigma-Aldrich unless stated.

7.1.2 General Procedures

7.1.2.1 Synthesis of Orthogonally Protected Heteroaromatic Amino Acids via Pd-Catalysed Negishi Cross-Coupling.

Reaction Conditions A:

Acid washed zinc dust (4.00 eq.) was heated under vacuum at 100 °C for 30 min whilst vigorously stirring. The zinc dust was cooled to 70 °C and placed under a positive pressure of argon, anhydrous DMF (0.5 mL) and I_2 (0.015 g) were added and the light grey suspension stirred for a further 20 min. The reaction mixture was cooled to 50 °C at which point iodoalanine (**1** or **9**) (1.00 eq.) was dissolved in DMF (0.5 mL) and added. Stirring under argon was continued for another 20 min followed by addition of the halo-aromatic (1.00 eq.), Tris(dibenzylideneacetone)dipalladium ($Pd_2(dba)_3$) (0.03 eq.) and tri(*o*-tolyl)phosphine (**P(*o*-tol)$_3$**) (0.10 eq.). The reaction mixture was left to stir under argon at 50 °C for 5 h, then overnight at room temperature. The cooled reaction

mixture was purified *via* column chromatography without work-up (SiO_2; 80/20 hexane/EtOAc → 100% EtOAc and if impure a second column was run using SiO_2; 100% CH_2Cl_2).

Reaction Conditions B:

This follows the same reaction procedure as above: procedure **A**, however, 2-dicyclohexylphosphino-2',6'-dimethoxybiphenyl (SPhos) is used as a replacement phosphine ligand to **P(*o*-tol)$_3$** in the same molar equivalents. The reaction mixture was purified as outlined in r*eaction conditions A*:

Reaction Conditions C:

Acid washed zinc dust (4.00 eq.) was heated under vacuum at 100 °C for 30 min whilst vigorously stirring. The reaction mixture was cooled to 50 °C and placed under a positive pressure of argon, anhydrous DMF (0.5 mL) and I_2 (0.015 g) were added and the light grey suspension stirred for a further 20 min. Further cooling of the reaction mixture to room temperature at which point iodoalanine (**7** or **20**) (1.00 eq.) was dissolved in DMF (0.5 mL) and added. Stirring under argon was continued for another twenty mins and the halo-aromatic (1.00 eq.), $Pd_2(dba)_3$ (0.03 eq.) and **SPhos** (0.10 eq.) were added. The reaction mixture was left to stir overnight at room temperature under an argon atmosphere. The reaction mixture was purified as outlined in *reaction conditions A*:

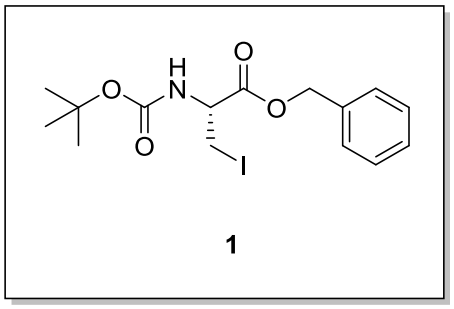
7.1.2.2 Solid Phase Cyclic Dipeptide (DKP) Synthesis

To pre-loaded MHBA-linker resin (**65**) (1.00 eq.) a pre-mixed peptide coupling solution of Boc protected amino acid (***A***) (3.00 eq.), DMAP (0.05 eq.) and DIC (3.00 eq) in DCM (2 mL) was added. The reaction mixture was shaken for 30 min at room temperature. The solution was drained and coupling procedure was repeated. The resin was washed with DCM (5 x 5 mL). A deprotection solution of 40% TFA in DCM (3 mL) was added to the resin and stirred for 10 min at room temperature. The solution was drained and this deprotection procedure was repeated. The resin was washed with DCM (5 x 5 mL). A pre-mixed peptide coupling solution of Boc-protected amino acid (***B***) (3.00 eq.), DIPEA (5.00eq.) and PyBOP (3.00 eq.) in DCM (3 mL) was added to the resin and stirred for 1 h at room temperature. The solution was drained and the coupling

procedure was repeated. The resin was washed with DCM (5 x 5 mL) a deprotection solution of 40% TFA in DCM (3 mL) was added to the resin and stirred for 10 min at room temperature. The solution was drained and deprotection procedure was repeated. The resin was washed with DCM (5 x 5 mL). Finally, 10% DIPEA in DCM (3 mL) was added to the resin and stirred for 10 min at room temperature. The filtrate solution was drained, collected and then repeated twice. The combined filtrate solutions were evaporated under reduced pressure to yield crude cyclic dipeptide. Dipeptides were purified *via* preparative TLC (SiO_2; 80/20% EtOAc/hexane) to yield pure cyclic dipeptide.

7.1.3 Small Molecule Syntheses

Benzyl 2-(*R*)-[(*tert*-butoxycarbonyl)amino]-3-iodopropanoate, Boc-β-iodo-Ala-OBzl. **1**

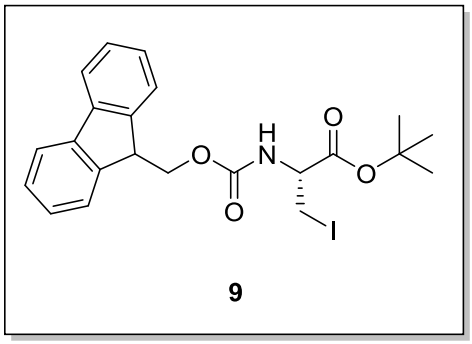


1

Boc-L-Ser-OBzl **10** (0.70 g, 2.20 mmol), I_2 (0.60 g, 2.20 mmol), imidazole (0.16 g, 2.20 mmol) and PPh_3 (0.62 g, 2.20 mmol) were dissolved in DCM (25 mL) and stirred for 20 h at room temperature. The reaction mixture was evaporated under reduced pressure and purified *via* column chromatography (SiO2; 95/5 % hexane/EtOAc → 40/60 % hexane/EtOAc) to yield the product **1** as pale yellow powder (0.71 g, 73%). δ_H (400 MHz; $CDCl_3$) 1.45 (9H, s, *t*Bu), 3.58(2H, m, ß-CH_2), 4.56 (1H, m, a-CH), 5.21 (2H, m, CH_2Ar), 5.25 (1H, d, *J* 7.6, NH) 7.33 to 7.48 (5H, m, ArH); *m/z* (ESI^+) 427.9 $[M+Na]^+$.

Data obtained is consistent with that given in the literature: $[\alpha]^D_{25}$ = -22.1° (c 1.0, $CHCl_3$), δ_H (300 MHz; $CDCl_3$) 1.45 (9H, s), 3.55 (2H, dd, *J* 3.6, 10.2), 4.55 (1H, m), 5.18 (1H, d, *J* 12.3), 5.24 (1H, d, *J* 12.3), 5.24 (1H, br d, *J* 7.5), 7.34 to 7.39 (5H, m); *m/z* (EI-MS^+) 270.0 $[M - CO_2Bn]^+$.[1]

(2*R*)-2-(9*H*-Fluoren-9-ylmethoxycarbonylamino)-3-iodo-propyric acid *tert*-butyl ester, Fmoc-β-iodo-Ala-O*t*Bu. **9**

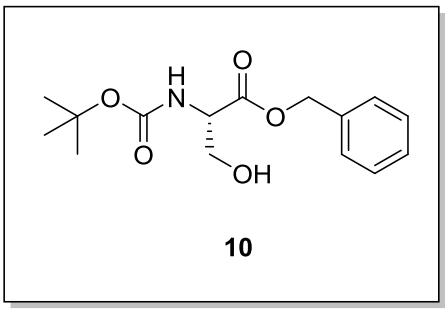


9

Fmoc-L-Ser-O*t*Bu **11** (1.64 g 4.30 mmol) dissolved in DCM (40 mL) and stirred at room temperature. Addition of triphenylphosphine (PPh_3) (1.12 g, 4.30 mmol), then imidazole (0.29 g, 4.30 mmol) followed by the addition of iodine (I_2) (1.09, 4.3 mmol). The reacting mixture was left to stir at room temperature for 20 h, evaporated under reduced pressure and purified *via* column chromatography (SiO_2; 90/10 hexane/EtOAc → 100% EtOAc) to yield product **9** as a sticky yellow solid (1.60 g, 77%). $[\alpha]_D^{24}$ = +15.36 (c 1.0 in $CHCl_3$); ν_{max}(solid)/cm^{-1} 3388 (w, N-H), 2968 (w, C-H), 1735 (s, C=O), 1698 (vs, C=O); δ_H (400 MHz; $CDCl_3$) 1.53 (9H, s, CO_2tBu), 3.61 (2H, m, ß-CH_2) 4.25 (1H, m, Fmoc-C**H**CH_2), 4.35 (1H, m, α-CH), 4.43 (2H, m, Fmoc-CHC**H**$_2$), 5.68 (1H, bd, *J* 6.8, NH), 7.31 to 7.35 (2H, m, Fmoc-ArH), 7.39 to 7.43 (2H, m, Fmoc-ArH), 7.60 (2H, bd, *J* 7.3, Fmoc-ArH), 7.77 (2H, bd, *J* 7.5, Fmoc-ArH); HRMS *m/z* (ES^+) 516.0648 ($[M+Na]^+$. $C_{22}H_{24}NO4INa$ requires 516.0625).

Data obtained is consistent with that given in the literature: $[\alpha]_D^{23}$ = +16.3 (c 1.1 in $CHCl_3$), δ_H ($CDCl_3$) 1.51 (9H, s), 3.58 (1H, dd, *J* 3.5, 10.5), 3.61 (1H, dd, *J* 3.5, 10.5), 4.24 (1H, t, *J* 7.5), 4.32 to 4.43 (3H, m), 5.72 (1H, d, *J* 6.5), 7.29 to 7.33 (2H, m), 7.39 (2H, t, *J* 7.5), 7.61 (2H, d, *J* 7.5), 7.75 (2H, d, *J* 7.5)*; m/z* (EI-MS^+) 493 ($[M+H]^+$).[2]

Benzyl 2-(*S*)-[(*tert*-butoxycarbonyl)amino]-3-hydroxypropanoate, Boc-Ser-OBzl. **10**

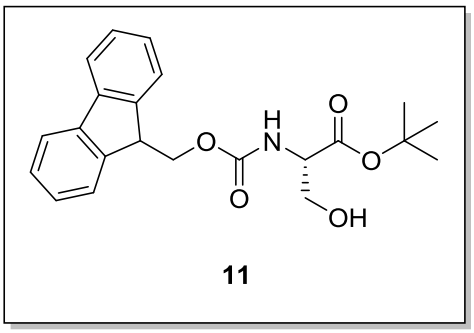


10

Boc-L-Ser-OH (1.53 g, 7.46 mmol), benzyl bromide (0.98 mL, 8.21 mmol) and K_2CO_3 (1.13 g, 8.21 mmol) were dissolved in DMF (30 mL) and stirred for 20 h at room temperature. H_2O (150 mL) was added and the aqueous phase was extracted with diethyl ether (3 x 60 mL). The combined organic extracts were washed with brine, (2 x 50 mL), dried over $MgSO_4$ and evaporated under reduced pressure to yield **10** as white powder (1.71 g, 78%). δ_H (400 MHz; $CDCl_3$) 1.44 (9H, s, *t*Bu), 3.95 (2H, m, ß-CH_2), 4.43 (1H, m, a-CH), 5.21 (2H, m, CH_2Ar), 5.48 (1H, m, NH) 7.29 to 7.41 (5H, m, ArH); *m/z* (ESI^+) 318.1 $[M+Na]^+$.

Data obtained is consistent with that given in the literature: δ_H (300 MHz; $CDCl_3$) 1.44 (9H, s), 2.41 (1H, t, *J* 6.1) 3.90 to 3.98 (2H, m), 4.41 (1H, bs), 5.21 (2H, s), 5.47 (1H, bs) 7.35 (5H, bs).[3]

(2*S*)-2-(9*H*-Fluoren-9-ylmethoxycarbonylamino)-3-hydroxy-propyric acid *tert*-butyl ester**,** Fmoc-Ser-O*t*Bu. **11**



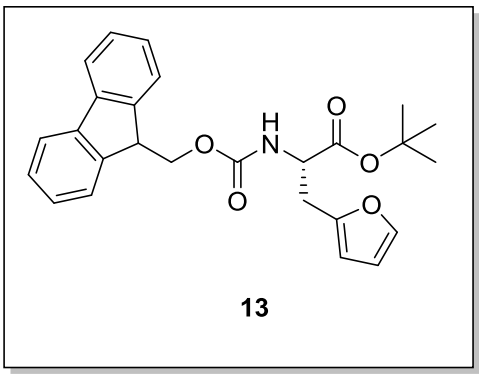
11

Fmoc-L-Ser-OH (2.00 g, 6.10 mmol) is dissolved in EtOAc (15 mL). *tert*-Butyl 2,2,2 – trichloroacetimidate (TBTA) (2.08 g, 9.50 mmol) was dissolved in EtOAc (20 mL) and added. The reaction was allowed to stir at room temperature for 20 h. The reaction mixture was evaporated under reduced pressure and purified *via* column chromatography (SiO_2; 80/20 hexane/EtOAc → 100% EtOAc). Further work-up to increase product purity involved re-dissolving in ether and washing with water (2 x 10 mL). The organic layer is dried with $MgSO_4$, filtered and evaporated to yield **11** as a white crystalline powder (1.53 g, 63%). ν_{max}(solid)/cm^{-1} 3374 (w bd, OH), 2963 (w, C-H), 1737 (s, C=O), 1679 (vs, C=O); δ_H (400 MHz; $CDCl_3$) 1.49 (9H, s, CO_2*t*Bu), 3.93 (2H, m, ß-CH_2) 4.22 (1H, m, Fmoc-C**H**CH_2), 4.32 (1H, m, α-CH), 4.43 (2H, d, *J* 7.6, Fmoc-CHC**H**$_2$), 5.68 (1H, bs, NH), 7.30 to 7.33 (2H, app. t, Fmoc-ArH), 7.38 to 7.42 (2H, app. t, Fmoc-ArH),

7.60 (2H, bd, *J* 7.2, Fmoc-ArH), 7.77 (2H, bd, *J* 7.3, Fmoc-ArH); HRMS *m/z* (ES^+) 406.1630 ($[M+Na]^+$. $C_{22}H_{25}NO_5Na$ requires 406.1612).

Data obtained is consistent with that given in the literature: $[\alpha]_D$ = +5.9 (c 1.2 in $CHCl_3$), δ_H ($CDCl_3$); δ_H (400 MHz; $CDCl_3$) 1.49 (9H, s), 3.93 (2H, bs), 4.22 (1H, t, *J* 6.8), 4.33 (1H, m), 4.41 (2H, d, *J* 6.8), 5.80 (1H, d, *J* 6.4, NH), 7.31 (2H, t, *J* 7.5), 7.40 (2H, t, *J* 7.5), 7.60 (2H, d, *J* 7.1), 7.76 (2H, d, *J* 7.5); *m/z* (ES^+) 406.1619 ($[M+Na]^+$).[4]

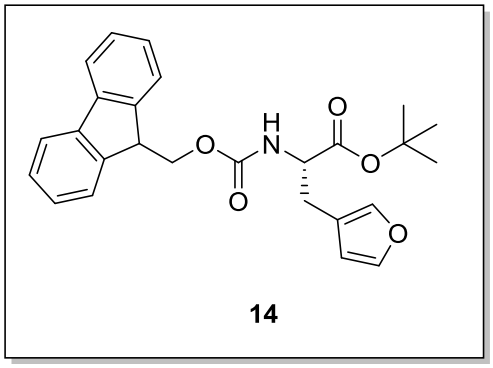
Fmoc-β-(2-furyl)-Ala-O*t*Bu. **13**



13

The general procedure for Pd-catalysed Negishi cross-coupling *reaction conditions A* was followed (0.61 mmol of **3**). **13** obtained as a yellow crystalline solid, 11% (0.07 mmol). ν_{max}(solid)/cm^{-1} 2928 (w, C-H), 1702 (s bd, C=O); $[\alpha]^D_{27}$ = -5.6° (c 1.0, $CHCl_3$); δ_H (400 MHz; $CDCl_3$) 1.47 (9H, s, $CO_2{}^tBu$), 3.20 (2H, d, *J* 5.2, β-CH_2), 4.24 (1H, app. t, Fmoc-C**H**CH_2), 4.38 (2H, m, Fmoc-CHC**H**$_2$), 4.55 (1H, dt, *J* 5.2, 8.0, α-CH), 5.50 (1H, bd, *J* 8.0, NH), 6.09 (1H, d, *J* 2.8, furan-CH), 6.31 (1H, m , furan-CH), 7.30 to 7.34 (2H, m, Fmoc-ArH), 7.32 (1H, m, furan-CH), 7.39 to 7.43 (2H, m, Fmoc-ArH), 7.60 (2H, bd, *J* 7.4, Fmoc-ArH), 7.77 (2H, d, *J* 7.4, Fmoc-ArH); δ_C (100 MHz; $CDCl_3$) 170.3, 155.7, 150.7, 144.0, 142.1, 141.4, 127.8, 127.2, 125.3, 120.1, 110.5, 108.1, 82.6, 67.2, 53.6, 47.3, 31.2, 28.0; HRMS *m/z* (ES^+) 456.1784 ($[M+Na]^+$. $C_{26}H_{27}NO_5Na$ requires 456.1787). General procedure for Pd-catalysed Negishi cross-coupling *reaction conditions B* was followed; Yellow crystalline solid, 42% (0.26 mmol).

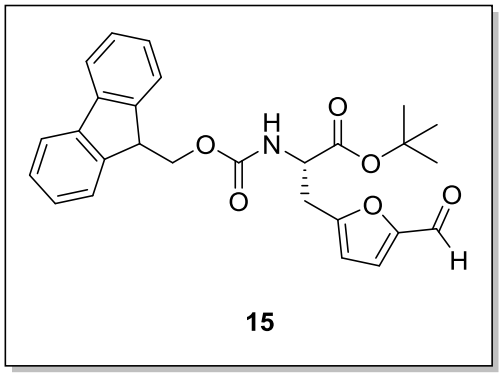
Fmoc-β-(3-furyl)-Ala-O*t*Bu. **14**



14

The general procedure for Pd-catalysed Negishi cross-coupling *reaction conditions B* was followed (0.61 mmol of **4**). **14** obtained as colourless oil, 27% (0.17 mmol). ν_{max}(solid)/cm^{-1} 3342 (w bd, N-H), 2979 (w, C-H), 1706 (s bd, C=O); $[\alpha]^{D}_{28}$ = -6.8° (c 1.0, $CHCl_3$); δ_H (400 MHz; $CDCl_3$) 1.46 (9H, s, $CO_2{}^tBu$), 2.95 (2H, m, β-CH_2), 4.43 (1H, m, Fmoc-C**H**CH_2), 4.42 (2H, m, Fmoc-CHC**H**$_2$), 4.51 (1H, m, α-CH), 5.35 (1H, bs, NH), 6.24 (1H, m, furan-CH), 7.23 (1H, m, furan-CH), 7.30 to 7.34 (2H, m, Fmoc-ArH), 7.37 (1H, m, furan-CH), 7.39 to 7.43 (2H, m, Fmoc-ArH), 7.57 to 7.60 (2H, m, Fmoc-ArH), 7.78 (2H, m, Fmoc-ArH); δ_C (100 MHz; $CDCl_3$) 170.6, 155.8, 144.1, 143.1, 141.5, 140.6, 127.9, 127.2, 125.2, 120.2, 119.2, 111.5, 82.6, 67.1, 54.4, 47.4, 31.0, 28.2; HRMS *m/z* (ES^+) 456.1785 ($[M+Na]^+$. $C_{26}H_{27}NO_5Na$ requires 456.1787).

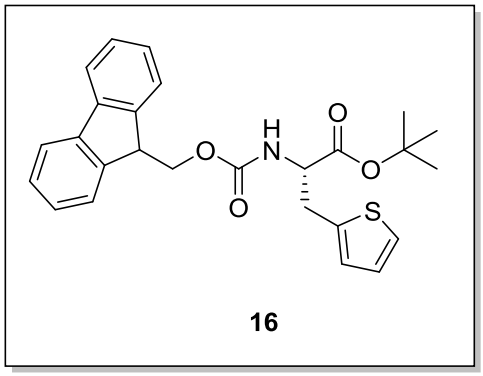
Fmoc-ß-(2-furfural)-Ala-O*t*Bu. **15**



15

The general procedure for Pd-catalysed Negishi cross-coupling *reaction conditions A* was followed (0.61 mmol of **8**). **15** obtained as a green oil, 46% (0.28 mmol). ν_{max}(solid)/cm^{-1} 3333 (w bd, N-H), 2979 (w, C-H), 1714 (s, C=O), 1677 (s, C=O); δ_H (400 MHz; $CDCl_3$) 1.47 (9H, s, CO_2tBu), 3.25 (1H, dd, *J* 5.6, 15.2, β-CH_2), 3.30 (1H, dd, *J* 5.6, 15.2, β-CH_2), 4.22 (1H, t, *J* 6.8, Fmoc-C**H**CH_2), 4.36 (1H, dd, *J* 6.8, 10, Fmoc-CHC**H**$_2$), 4.43 (1H, dd, *J* 6.8, 10, Fmoc-CHC**H**$_2$), 4.57 (1H, dd, *J* 5.6, 7.6, α-CH), 5.51 (1H, bd, *J* 7.6, NH), 6.30 (1H, d, *J* 3.6 furan-CH), 7.15 (1H, d, *J* 3.6, furan-CH), 7.27 to 7.31 (2H, m, Fmoc-ArH), 7.36 to 7.40 (2H, m, Fmoc-ArH), 7.55 to 7.58 (2H, app. t, Fmoc-ArH), 7.77 (2H, bd, *J* 7.6, Fmoc-ArH), 9.55 (1H, s, furan-COH); δ_C (100 MHz; $CDCl_3$) 188.5, 177.2, 170.0, 157.9, 155.7, 152.7, 143.9, 141.5, 127.9, 127.2, 125.2, 120.2, 111.3, 83.4, 67.2, 53.2, 47.3, 31.7, 28.0; HRMS *m/z* (ES^+) 484.1746 ($[M+Na]^+$. $C_{27}H_{27}NO_6Na$ requires 484.1736).

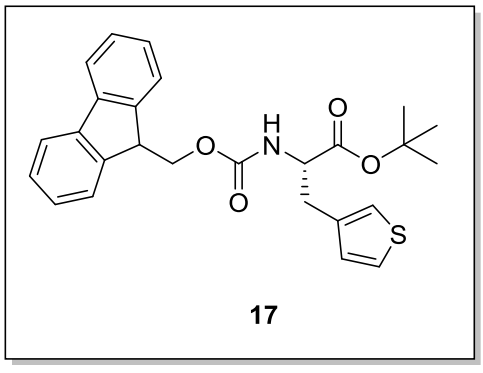
Fmoc-β-(2-thienyl)-Ala-O*t*Bu. **16**



16

The general procedure for Pd-catalysed Negishi cross-coupling *reaction conditions B* followed (0.62 mmol of **7**). **16** obtained as a yellow oil, 26% (0.16 mmol). ν_{max}(solid)/cm^{-1} 3342 (w bd, N-H), 2977 (w, C-H), 1710 (s bd, C=O); $[\alpha]^{D}_{27}$ = +15.0° (c 1.0, $CHCl_3$); δ_H (400 MHz; $CDCl_3$) 1.48 (9H, s, $CO_2{}^tBu$), 3.39 (2H, m, β-CH_2) 4.26 (1H, t, *J* 7.2, Fmoc-C**H**CH_2), 4.35 (1H, dd, 7.2, 10.4, Fmoc-CHC**H**$_2$), 4.48 (1H, dd, 7.2, 10.4, Fmoc-CHC**H**$_2$), 4.58 (1H, m, α-CH), 5.90 (1H, bd, *J* 8.0, NH), 6.83 (1H, d, *J* 3.0 thiophene-CH), 6.96 (1H, dd, *J* 3.0, 5.0, thiophene-CH), 7.20 (1H, d, *J* 5.0, thiophene-CH), 7.32 to 7.36 (2H, m, Fmoc-ArH), 7.40 to 7.44 (2H, m, Fmoc-ArH), 7.61 to 7.64 (2H, m, Fmoc-ArH), 7.78 (2H, bd, *J* 7.6, Fmoc-ArH); δ_C (100 MHz; $CDCl_3$) 170.0, 155.7, 144.0, 141.4, 137.6, 127.8, 127.1, 127.0, 126.9, 125.2, 124.8, 120.1, 83.8, 67.1, 55.0, 47.3, 32.5, 28.1; HRMS *m/z* (ES^+) 472.1562 ($[M+Na]^+$. $C_{26}H_{27}NO_4SNa$ requires 472.1559). General procedure for Pd-catalysed Negishi cross-coupling *reaction conditions B* followed with **5**; Yellow oil, 26% (0.15 mmol). General procedure for Pd-catalysed Negishi cross-coupling *reaction conditions C* followed with **7**; Yellow oil, 22% (0.13 mmol).

Fmoc-β-(3-thienyl)-Ala-O*t*Bu. **17**

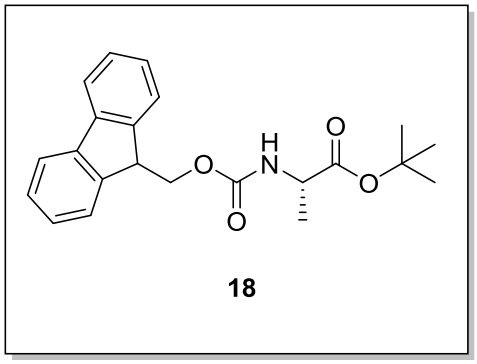


17

The general procedure for Pd-catalysed Negishi cross-coupling *reaction conditions B* followed (0.61 mmol of **6**). **16** obtained as a yellow oil, 20% (0.12 mmol). ν_{max}(solid)/cm^{-1} 3342 (w bd, N-H), 2976 (w, C-H), 1707 (s bd, C=O); $[\alpha]^{D}_{28}$ = -18.1° (c 1.0, $CHCl_3$); δ_H (400 MHz; $CDCl_3$) 1.45 (9H, s, CO_2tBu), 3.15 (2H, m, β-CH_2) 4.23 (1H, m, Fmoc-C**H**CH_2), 4.30 to 4.57 (3H, m, Fmoc-CHC**H**$_2$, α-CH), 5.35 (1H, bs, NH), 6.92 (1H, d, *J* 4.0, thiophene-CH), 6.98 (1H, m, thiophene-CH), 7.26 (1H, m, thiophene-CH), 7.29 to 7.35 (2H, m, Fmoc-ArH), 7.39 to 7.44 (2H, m, Fmoc-ArH), 7.58 - 7.61 (2H, m Fmoc-ArH) 7.77 (2H, d, *J* 7.6, Fmoc-ArH); δ_C (100 MHz; $CDCl_3$) 170.7,

155.7, 144.0, 141.4, 136.3, 128.7, 127.8, 127.2, 125.8, 125.2, 122.9, 120.1, 82.5, 67.0, 54.8, 47.3, 32.9, 28.0; HRMS *m/z* (ES^+) 472.1563 ($[M+Na]^+$. $C_{26}H_{27}NO_4SNa$ requires 472.1559).

(2S)-2-(9*H*-Fluoren-9-ylmethoxycarbonylamino)propionic acid tert-butyl ester, Fmoc-Ala-O*t*Bu. **18**

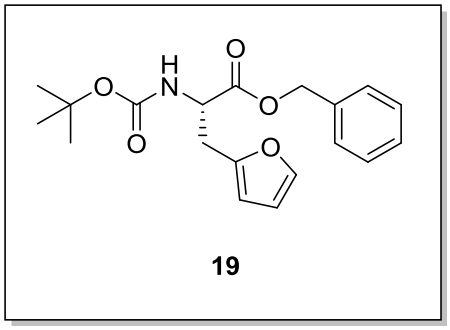


18

The general procedure for Pd-catalysed Negishi cross-coupling *reaction conditions A* followed (0.61 mmol). **18** was purified from desired product and obtained as a pale yellow crystalline solid, 82% (0.49 mmol). δ_H (400 MHz; $CDCl_3$) 1.40 (3H, d, *J* 7.2, β-CH_3), 1.48 (9H, s, CO_2*t*Bu), 4.19 to 4.29 (1H, m, Fmoc-C**H**CH_2) 4.24 (1H, m, α-CH), 4.39 (2H, d, *J* 7.4, Fmoc-CHC**H**$_2$), 5.35 (1H, bd, *J* 6.8, NH), 7.31 to 7.35 (2H, app. t, Fmoc-ArH), 7.39 to 7.43 (2H, app. t, Fmoc-ArH), 7.60 (2H bd, *J* 7.6, Fmoc-ArH), 7.77 (2H, bd, *J* 7.6, Fmoc-ArH); δ_C (100 MHz; $CDCl_3$) 170.2, 155.7, 141.5, 127.8, 127.2, 125.3, 125.2, 120.1, 82.2, 67.1, 50.4, 47.4, 28.2, 18.7; *m/z* (ES^+) 390.3 $[M+Na]^+$.

Data obtained is consistent with that given in the literature: δ_H (400 MHz; $CDCl_3$) 1.40 (3H, d, *J* 6.6), 1.48, (9H, s), 4.20 to 4.30 (2H, m), 4.38 (2H, d, *J* 7.4), 5.37 (1H, d, *J* 6.6); 7.32 (2H, td, *J* 1.5, 7.4), 7.40 (2H, t, *J* 7.4); 7.61 (2H, d, *J* 7.4), 7.77 (2H, d, *J* 7.4).[5]

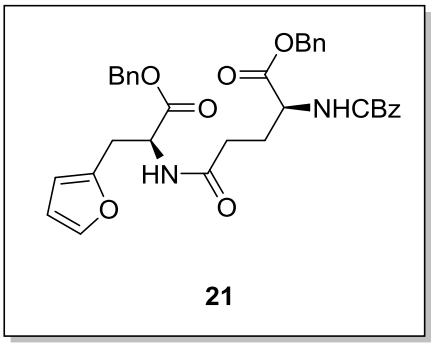
Boc-β-(2-furyl)-Ala-OBzl. **19**



19

The general procedure for Pd-catalysed Negishi cross-coupling *reaction conditions B* followed (0.65 mmol of **3**). **19** was obtained as a yellow oil, 86% (0.56 mmol). $[\alpha]^{D}_{28}$ -7.0° (c 1.0, CH_3Cl); δ_H (400 MHz; $CDCl_3$) 1.43 (9H, s, *t*Bu), 3.14 (2H, m, β-CH_2), 4.61 (1H, m, α-CH), 5.46 (2H, m, CH_2Ar), 5.95 (1H, d, *J* 3.2, furan-CH), 6.24 (1H, dd, *J* 2.0, 3.2, furan-CH), 7.27 (1H, d, *J*, 2.0, furan-CH), 7.31 to 7.40 (5H, m, ArH); δ_C (100 MHz; $CDCl_3$) 171.4, 155.2, 150.4, 142.2, 135.4, 128.6, 128.5, 110.4, 108.0, 80.1, 67.3, 52.9, 31.0 28.4, 18.6; HRMS *m/z* (ES^+) 368.1473 ($[M+Na]^+$. $C_{19}H_{23}NO_5Na$ requires 368.1474).

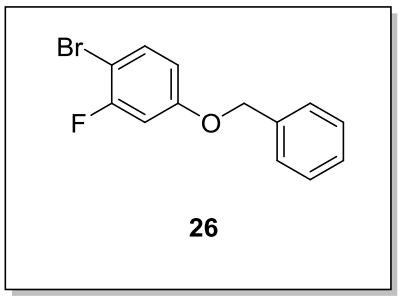
Z-Glu-(-β-(2-Furyl)-Ala-OBn)-OBn. **21**



21

Boc-β-(2-furyl)-Ala-OBn **19** (0.05 g, 0.14 mmol) was stirred in TFA: DCM 1:4 (10 mL) until the Boc deprotection was seen to be complete by TLC. The reaction mixture was then evaporated under reduced pressure to dryness. The resulting yellow powder was dissolved in DCM (15 mL). Z-Glu-OBn (0.05 g, 0.14 mmol), PyBOP (0.08 g, 0.14 mmol) DIPEA (0.08 mL, 0.42 mmol) were added and the reaction mixture was stirred overnight at room temperature. The reaction mixture was evaporated under reduced pressure and purified *via* column chromatography (SiO_2; 95/5% hexane/EtOAc → 40/60% hexane/EtOAc) to yield **21** as a white powder (0.07 g, 78%). δ_H (400 MHz; $CDCl_3$) 1.90 to 2.34 (4H, m, β-CH and γ-CH_2), 3.18 (2H, m, β-CH_2), 4.46 (1H, m, α-CH) 4.87 (1H, m, α-CH), 5.08 to 5.28 (6H, m, CH_2Ar), 5.62 (1H, d, *J* 8.0, NH), 5.94 (1H, d, *J* 2.8, furan-CH) 6.19 (1H, t, *J* 2.8, furan-CH), 6.28 (1H, d, *J* 7.2, NH), 7.21 (1H, m, furan-CH), 7.27 to 7.41 (15H, m, ArH); δ_C (100 MHz; $CDCl_3$) 171.0, 150.3, 142.3, 136.3. 135.3, 128.8, 128.7, 128.6, 128.5, 128.4, 128.3, 128.2, 110.4, 108.2, 67.6, 67.5, 67.2, 53.6, 51.8, 48.4, 32.2, 30.6, 28.4, 18.5; HRMS *m/z* (ES^+) 599.2387 ($[M+H]^+$. $C_{34}H_{34}N_2O_8$ requires 599.2393).

4-(Benzyloxy)-1-bromo-2-fluorobenzene. **26**

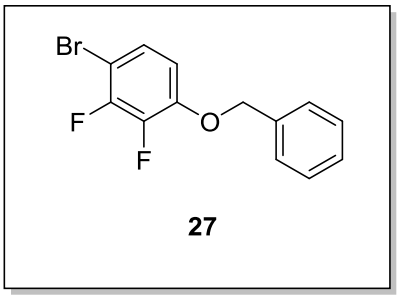


26

4-Bromo-3-fluorophenol **24** (0.33 g, 1.74 mmol), benzyl bromide (0.23 mL, 1.91 mmol) and K_2CO_3 (0.29 g, 1.91 mmol) was stirred in MeCN (25 mL) for 20 h at room temperature. The reaction mixture was evaporated under reduced pressure and purified *via* column chromatography (SiO_2; 95/5 % hexane/EtOAc → 40/60 % hexane/EtOAc) to yield the product as colourless oil (0.276 g, 58%). δ_H (400 MHz; $CDCl_3$) 5.04 (2H, s, CH_2Ar), 6.68 (1H, m, ArH), 6.78 (1H, m, ArH), 7.30 to 7.45 (6H, m, ArH); δ_F (376 MHz; $CDCl_3$) -105.14 (1F, t, *J* 9.2, ArF); *m/z* (ES^-) 278.9 $[M-H]^-$, ^{79}Br.

Data obtained is consistent with that given in the literature: δ_H (400 MHz; $CDCl_3$) 5.05 (2H, s), 6.69 (1H, ddd, *J* 4.0, 8.0, 10.0), 6.77 (1H, dd, *J* 40, 8.0), 7.32 to 7.42 (6H, m).[6]

4-(Benzyloxy)-1-bromo-2,3-difluorobenzene. **27**

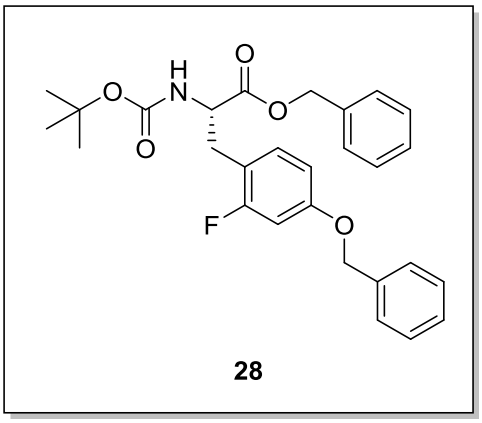


27

4-Bromo-2,3-difluorophenol **25** (0.31 g, 1.49 mmol), benzyl bromide (0.20 mL, 1.64 mmol) and K_2CO_3 (0.25 g 1.64 mmol) was stirred in MeCN (25 mL) for 20 h at room temperature. The reaction mixture was evaporated under reduced pressure and purified via column chromatography (SiO_2; 95/5 % hexane/EtOAc → 40/60 % hexane/EtOAc) to yield the product as a colourless oil (0.350 g, 77%). δ_H (400 MHz; CDCl3) 5.18 (2H, s, CH_2Ar), 6.85 (1H, m, ArH), 7.27 (1H, m, ArH), 7.38 (3H, m, ArH), 7.52 (2H, m, ArH); δ_F (376 MHz; CDCl3) -136.1 (1F, m, ArF), -149.4 (1F, m, ArF); *m/z* (ESI^-) 296.9 $[M-H]^-$, ^{79}Br.

Data obtained is consistent with that given in the literature: δ_H (400 MHz; $CDCl_3$) 5.16 (2H, s), 6.90 (1H, dt, *J* 1.8, 8.0), 7.19 (1H, dt, *J* 2.5, 8.1), 7.42 (5H, m).[7]

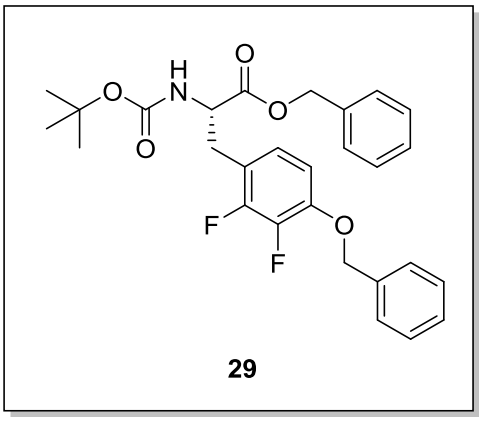
Boc-2-fluoro-Tyr(OBzl)-OBzl. **28**



28

The general procedure for Pd-catalysed Negishi cross-coupling *reaction conditions B* followed (0.72 mmol of **26**). The reaction mixture was purified *via* column chromatography (SiO_2; 98/2 % hexane/EtOAc → 50/50 % hexane/ EtOAc) to yield **28** as a yellow powder, 72% (0.52 mmol). ν_{max}(solid)/cm^{-1} 3364 (w bd, N-H), 2976 (w, C-H), 1732 (s, C=O) 1687 (s, C=O); $[\alpha]^{D}_{29}$ = -11.8° (c 1.0, $CHCl_3$); δ_H (400 MHz; CDCl3) 1.41 (9H, s, *t*Bu), 3.10 (2H, m, ß-CH_2), 4.58 (1H, m, α-CH), 5.00 (2H, s, CH_2Ar), 5.05 (1H, m, NH), 5.14 (2H, s, CH_2Ar), 6.62 to 6.68 (2H, m, ArH), 6.94 (1H, t, *J* 9.0, ArH), 7.30 to 7.44 (10H, m, ArH); δ_F (376 MHz; $CDCl_3$) -115.0 (1F, m, ArF); δ_C (150 MHz; $CDCl_3$) 171.8, 161.9 (1C, d, *J* 208.5), 159.3 (m), 155.2, 136.6, 135.4, 128.8, 128.7, 128.6, 128.3, 128.6, 127.6, 132.1 (1C, d, *J* 4.5), 115.2 (1C, d, *J* 15.0), 110.9 (1C, d, *J* 3.0), 102.6 (1C, d, *J* 22.5), 80.0, 70.5, 67.4, 54.0, 28.4, 18.8; HRMS *m/z* (ES^+) 502.1994 ([M+ Na]$^+$. $C_{28}H_{30}NO_5F$ requires 502.2006).

Boc-2,3-difluoro-Tyr(OBzl)-OBzl. **29**

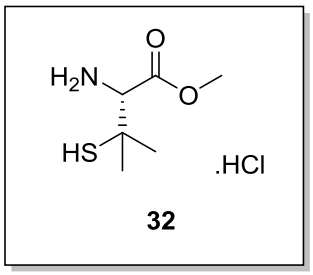


29

The general procedure for Pd-catalysed Negishi cross-coupling *reaction conditions B* followed (0.72 mmol of **27**). The reaction mixture was purified *via* column chromatography (SiO_2; 98/2 % hexane/EtOAc → 50/50 % hexane/EtOAc) to yield **29** as brown oil, 60% (0.43 mmol). ν_{max}(solid)/cm^{-1} 3377 (w bd, N-H), 2965 (w, C-H), 1737 (s, C=O), 1686 (s, C=O); $[\alpha]^{D}_{29}$ = -19.4° (c 1.0, $CHCl_3$); δ_H (400 MHz; $CDCl_3$) 1.40 (9H, s, *t*Bu), 2.96 (2H, m, ß-CH_2), 4.55 (1H, m, α-CH), 5.05 to 5.24 (4H, m, 2 x CH_2Ar), 5.06 (1H, m, NH), 6.72 to 6.80 (2H, m, ArH), 7.23 to 7.45 (10H, m, ArH); δ_F (376 MHz; CDCl3) -137.5 (1F, m, ArF), -153.0 (1F, m, ArF); δ_C (150 MHz; $CDCl_3$) 171.8, 155.2, 150.8 (1C, dd, *J* 9.0, 211.5),

146.1 (1C, d, *J* 6.0), 144.3 (1C, dd, *J* 11.3, 213.0), 136.5, 135.4, 128.7, 128.7, 128.7, 128.6, 128.5, 128.4, 126.2 (br s), 124.9 (m), 111.2 (1C, d, *J* 13.5), 80.0, 75.8, 67.2, 54.1, 28.4, 18.8; HRMS *m/z* (ES^+) 520.1921 ($[M+Na]^+$. $C_{28}H_{29}NO_5F_2$ requires 520.1912).

H-Pen-OMe. **32**

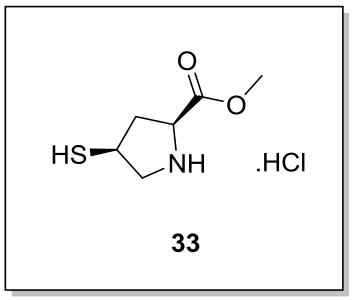


32

Penicillamine (0.30 g, 2.01 mmol) was dissolved in MeOH (20 mL) and thionyl chloride was added (0.29 mL, 4.02 mmol) stirred at reflux for 15 h. Additional thionyl chloride was added (0.29 mL, 4.02 mmol) stirred at reflux for 40 h. Solvent was evaporated under reduced pressure to yield **32** as a pale yellow powder (quant.). δ_H (400 MHz; d_4-methanol); 1.48 (3H, s, β-CH_3), 1.55 (3H, s, β-CH_3), 3.86 (3H, s, OCH_3), 4.12 (1H, s, α-CH). δ_C (376 MHz; d_4-methanol) 167.3, 62.5, 42.9, 29.0, 26.2; *m/z* (ES^+) 164.1 $[M+H]^+$.

Data obtained is consistent with that given in the literature: δ_H (400 MHz; d_4-methanol) 1.47 (3H, s), 1.56 (3H, s), 3.86 (3H, s). 4.12 (1H, s).[8]

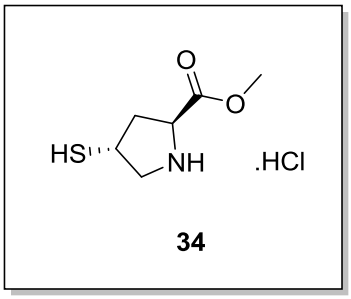
Cis-4-(thio)-L-proline methyl ester. **33**



33

35 (0.05 g, 0.17 mmol) was dissolved in MeOH (5 mL). Thionyl chloride (0.04 mL, 0.50 mmol) was slowly syringed into the reaction vessel at 5°C and the reaction mixture was then heated to reflux for 4 h. The reaction mixture was evaporated under reduced pressure to yield **33** as a pale yellow powder (quant.). ν_{max}(solid)/cm^{-1} 2937 (w, C-H), 1702 (s, C=O), 1333; $[\alpha]^D_{27}$ = -5.1° (c 0.1, $CHCl_3$); δ_H (400 MHz; $CDCl_3$) 2.10 to 2.20 (2H, m, ß-CH, SH), 2.86 (1H, m, ß-CH), 3.42 (1H, m, δ-CH), 3.65 (1H, m, γ-CH), 3.82 to 3.93 (4H, m, OCH_3, δ-CH), 4.57 (1H, m. α-CH), 9.11 (1H, m, NH), 11.31 (1H, m, NH); δ_C (100 MHz; $CDCl_3$) 169.1, 58.8, 54.2, 54.0, 38.9, 34.9; HRMS *m/z* (ES^+) 162.0578 ($[M+H]^+$. $C_6H_{12}NO_2S$ requires 162.0589).

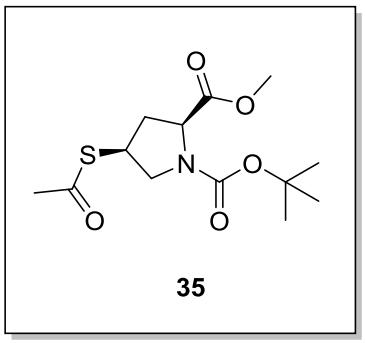
Trans-4-(thio)-L-proline methyl ester. **34**



34

36 (0.05 g, 0.17 mmol) was dissolved in MeOH (5 mL). Thionyl chloride (0.04 mL, 0.50 mmol) was slowly syringed into the reaction vessel at 5°C and the reaction mixture was then heated to reflux for 4 h. The reaction mixture was evaporated under reduced pressure to **34** a pale yellow powder (quant.). ν_{max}(solid)/cm^{-1} 3586, 2946 (w, C-H), 1734 (s, C=O), 1239; $[\alpha]^{D}_{27}$ = +6.2° (c 0.2, $CHCl_3$); δ_H (400 MHz; $CDCl_3$) 2.35 (1H, m, ß-CH), 2.48 (1H , d, *J* 8.0, SH) 2.67 (1H, m, ß-CH), 3.37 (1H, dd, *J* 7.2, 12.0, δ-CH), 3.60 (1H, m, γ-CH), 3.85 (3H, s, OCH_3), 4.05 (1H, dd, *J* 7.2, 12.0, δ-CH), 4.74 (1H, m. α-CH), 9.43 (1H, m, NH), 10.70 (1H, m, NH); δ_C (100 MHz; $CDCl_3$) 169.0, 58.8, 54.5, 53.9, 38.9, 34.8; HRMS *m/z* (ES$^+$) 162.0580 ([M+H]$^+$. $C_6H_{12}NO_2S$ requires 162.0589).

N-Boc-*cis*-4-(acetylthio)-L-proline methyl ester. **35**

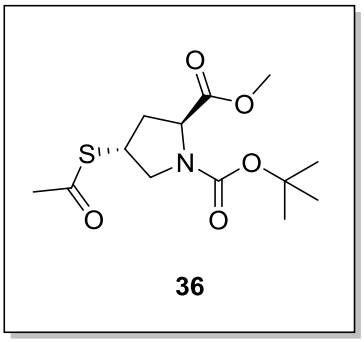


35

Potassium thioacetate (0.06 g, 0.50 mmol) was added to **43** (0.12 g, 0.33 mmol) dissolved in DMF (8 mL). The reaction mixture was stirred at 70 °C for 3 h. The reaction mixture was evaporated under reduced pressure and purified *via* column chromatography (SiO_2; 90/10 hexane/EtOAc → 100% EtOAc) yielding **35** as a colourless oil (quant.). δ_H mixture of two rotamers in a ratio of ca. 1.0:1.1, (400 MHz; $CDCl_3$) 1.37 and 1.43 (9H, 2 s, *t*Bu), 1.95 (1H, m, ß-CH), 2.31 (3H, s, CH_3CO), 2.70 (1H, m, ß-CH), 3.32 (1H, m, δ-CH), 3.72 (3H, OCH_3), 3.88 to 4.00 (2H, m, γ-CH, δ-CH), 4.27 and 4.25 (1H, 2 t, α-CH) δ_C (100 MHz; $CDCl_3$) 195.1, 194.8, 172.9, 172.7, 153.9, 153.3, 80.6, 58.7, 58.2, 52.5, 52.3, 52.2, 51.4, 39.5, 38.8, 37.1, 35.8, 30.6, 28.5, 28.3; *m/z* (ES$^+$) 326.0 [M+Na]$^+$.

Data obtained is consistent with that given in the literature: $[\alpha]^{D}_{20}$ = −43.3° (c 1.2, $CHCl_3$); δ_H (400 MHz; $CDCl_3$) 1.44 (9H, s), 1.98 (1H, m), 2.34 (3H, s), 2.75 (1H, m), 3.36 (1H, m), 3.76 (3H, s, $COOCH_3$), 3.99 (2H, m), 4.40 (1H, m).[9]

N-Boc-*trans*-4-(acetylthio)-L-proline methyl ester. **36**

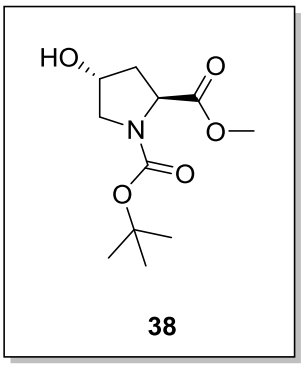


36

Potassium thioacetate (0.06 g, 0.55 mmol) was added to **42** (0.13 g, 0.37 mmol) dissolved in DMF (8 mL). The reaction mixture was stirred at 70 °C for 3 h. The reaction mixture was evaporated under reduced pressure and purified *via* column chromatography (SiO_2; 90/10% hexane/EtOAc → 100% EtOAc) yielding **36** as a colourless oil (0.110 g, 98%). δ_H mixture of two rotamers in a ratio of ca. 1.0:1.3 (400 MHz; $CDCl_3$) 1.41 and 1.45 (9 H, 2 s, *t*Bu), 2.16 to 2.44 (5H, m, CH_3CO, β-CH_2), 3.30 and 3.42 (1H, 2 m, δ-CH) 3.74 (3H, s, OCH_3) 3.93 (1H, m, δ-CH) 4.04 (1H, m, γ-CH) 4.31 and 4.39 (1H, 2 m, α-CH) δ_C (100 MHz; $CDCl_3$) 194.9, 194.8, 173.0, 172.7, 154.1, 153.4, 80.6, 58.6, 58.2, 52.5, 52.3, 52.1, 51.6, 39.6, 39.6, 37.0, 35.7, 30.7, 28.5, 28.3; HRMS *m/z* (ES^+) 326.1029 ($[M+Na]^+$. $C_{13}H_{21}NO_5SNa$ requires 326.1038).

Data obtained is consistent with that given in the literature: $[\alpha]^D_{20} = -33.0°$ (c 0.9, $CHCl_3$); δ_H (400 MHz; $CDCl_3$) 1.41 (9H, s), 2.10 (1H, m), 2.34 (3H, s), 2.47 (1H, m), 3.31 to 4.10 (3H, m), 3.75 (3H, s), 4.40 (1H, m).[9]

N-Boc-*trans*-4-(hydroxy)-L-proline methyl ester, Boc-Hyp-OMe. **38**

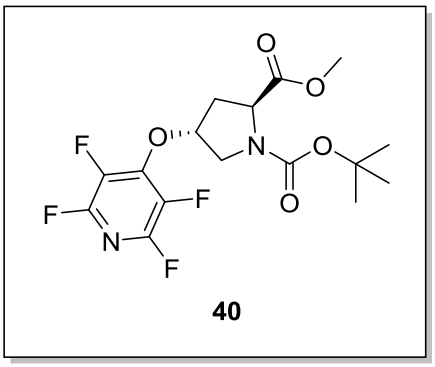


38

L-4-hydroxyproline methyl ester hydrochloride (2.88 g, 20.00 mmol), Boc anhydride (5.20 g, 23.80 mmol) and triethylamine (4.05 g, 40.0 mmol) were dissolved in DCM (30 mL) and stirred at room temperature for 24 h. The reaction mixture was evaporated under reduced pressure and purified *via* column chromatography (SiO_2; 100% hexane → 40/60% hexane/EtOAc) yielding **38** as a white crystalline solid (3.03 g, 62%). δ_H (400 MHz; $CDCl_3$) 1.37 (9H, s, *t*Bu), 2.04 (1H, m), 2.13 (2H, m, ß-CH_2), 3.49 (2H, m, δ-CH_2), 3.67 (3H, s, OCH_3), 4.28 to 4.48 (2H, m, α-CH, γ-CH); δ_C (100 MHz, $CDCl_3$) 173.8, 154.2, 80.5, 69.3, 58.1, 54.7, 52.3, 52.1, 39.1, 38.5, 28.3, 28.1; *m/z* (ES^+) 268.6 $[M+Na]^+$.

Data obtained is consistent with that given in the literature: δ_H (200 MHz; $CDCl_3$) 1.41 (9H, s), 2.06 (1H, m), 2.30 (2H, m), 3.58 to 3.62 (2H, m), 3.74 (3H, s), 4.40 to 4.49 (2H, m).[10]

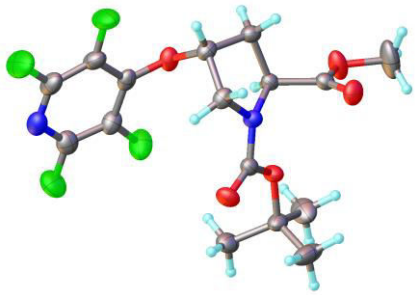
N-Boc-*trans*-4-(tetrafluoropyridyl)-L-proline methyl ester. **40**



40

Boc-Hyp-OMe **38** (0.23 g, 0.94 mmol), Cs_2CO_3 (0.26 g, 1.88 mmol) and pentafluoropyridine (0.31 mL, 2.82 mmol) were dissolved in MeCN (8 mL) and stirred at room temperature for 72 h. The reaction mixture was filtered and evaporated under reduced pressure. The crude product was purified *via* column chromatography (SiO_2; 90/10% hexane/EtOAc → 100% EtOAc) yielding **40** as a white powder (0.32 g, 85%); δ_H mixture of two rotamers in a ratio of ca. 1.0:2.0 (400 MHz; $CDCl_3$) 1.42 and 1.45 (9H, s, *t*Bu), 2.28 (1H, m, CH), 2.57 to 2.70 (1H, m, CH), 3.72 to 3.98 (2H, m, CH_2) 3.76 (3H, s, OCH_3) 4.43 (1H, m, CH), 4.36 (1H, m, CH); δ_F (376 MHz; $CDCl_3$) -89.5 (1F, m, ArF), 158.0 (1F, m, ArF); δ_C (100 MHz; $CDCl_3$) 172.9, 154.1, 153.7, 144.4 (2C, d, *J* 242.1), 135.3 (2C, d, *J* 293.2) 81.6, 57.8, 56. 9, 51.7, 37.5, 36.4, 28.4.

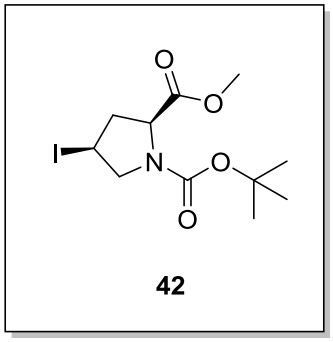
Crystal data and structure refinement:



Empirical formula	$C_{16}H_{18}N_2O_5F_4$
Formula weight	394.32
Temperature/K	120
Crystal system	orthorhombic
Space group	$P2_12_12_1$
a/Å	9.9914(5)
b/Å	10.0506(5)
c/Å	18.3781(9)
α/°	90

β/°	90
γ/°	90
Volume/Å^3	1845.52(16)
Z	4
ρ_{calc}g/cm^3	1.419
μ/mm^{-1}	1.151
F(000)	816.0
Crystal size/mm^3	0.387 × 0.113 × 0.103
Radiation	CuKα (λ = 1.54184)
2Θ range for data collection/°	9.626 to 150.38
Index ranges	-10 ≤ h ≤ 11, -12 ≤ k ≤ 11, -22 ≤ l ≤ 21
Reflections collected	22217
Independent reflections	3653 [R_{int} = 0.0286, R_{sigma} = 0.0166]
Data/restraints/parameters	3653/0/248
Goodness-of-fit on F^2	1.076
Final R indexes [I>=2σ (I)]	R_1 = 0.0325, wR_2 = 0.0835
Final R indexes [all data]	R_1 = 0.0347, wR_2 = 0.0854
Largest diff. peak/hole / e Å^{-3}	0.25/-0.20
Flack parameter	0.01(3)

N-Boc-*cis*-4-(iodo)-L-proline methyl ester, Boc-*cis*-iodoproline-OMe. **42**

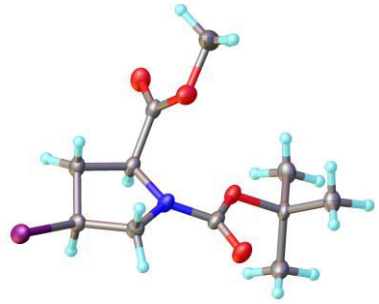


42

Boc-Hyp-OMe **38** (0.40 g, 1.64 mmol), imidazole (0.11 g, 1.64mmol), PPh_3 (0.43 g, 1.64 mmol) and I_2 (0.42 g, 1.64 mmol) were dissolved in DCM (20 mL) and stirred at room at room temperature for 4 h. The reaction mixture was evaporated under reduced pressure and purified *via* column chromatography (SiO_2; 100% hexane → 60/40% hexane/EtOAc) yielding **42** as a pale yellow crystalline solid (0.40 g, 69%). R_f 0.19 (hexane/EtOAc 9:1). δ_H mixture of two rotamers in a ratio of ca. 1.0:1.3 (400 MHz; $CDCl_3$) 1.40 and 1.46 (9H, 2 s, BocH), 2.34 (1H, m, ß-CH), 2.86 (1H, m, ß-CH), 3.66 (1H, dd, *J* 7.8, 10.2, δ-CH), 3.75 (3H, s, OCH_3), 4.00 to 4.13 (2H, m, γ-CH, δ-CH), 4.23 and 4.31 (1H, 2 t, *J* 7.8, α-CH); δ_C (100 MHz, $CDCl_3$) 172.3, 172.0, 153.4, 152.8, 80.9, 59.3, 58.7, 57.1, 56.8, 52.4, 43.0, 42.0, 28.5, 28.4, 12.9, 12.1; *m/z* (ES^+) 378.0 $[M+Na]^+$.

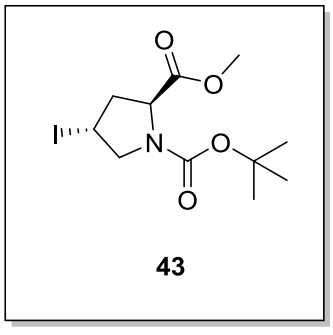
Data obtained is consistent with that given in the literature: $[\alpha]_D^{20}$ -20.9 (c 0.82°, $CHCl_3$); δ_H (500 MHz; $CDCl_3$) 1.38 and 1.43 (9H, 2 s), 2.26 to 2.35 (1H, m), 2.82 to 2.85 (1H, m), 3.63 (1H, dd, *J* 8.2, 10.2), 3.72 (3H, s), 4.00 to 4.10 (2H, m), 4.20 and 4.28 (1H, 2 t, *J* 7.5). [11]

Crystal data and structure refinement:



Empirical formula	$C_{11}H_{18}INO_4$
Formula weight	355.16
Temperature/K	120
Crystal system	monoclinic
Space group	$P2_1$
a/Å	5.5923(3)
b/Å	9.1371(6)
c/Å	13.7963(8)
α/°	90.00
β/°	99.189(2)
γ/°	90.00
Volume/Å^3	695.91(7)
Z	2
ρ_{calc}mg/mm^3	1.695
m/mm^{-1}	2.304
F(000)	352.0
Crystal size/mm^3	0.44 × 0.14 × 0.06
2Θ range for data collection	3 to 56°
Index ranges	$-7 \leq h \leq 7$, $-11 \leq k \leq 12$, $-18 \leq l \leq 18$
Reflections collected	7114
Independent reflections	3336[R(int) = 0.0219]
Data/restraints/parameters	3336/1/226
Goodness-of-fit on F^2	1.050
Final R indexes [I>=2σ (I)]	R_1 = 0.0262, wR_2 = 0.0615
Final R indexes [all data]	R_1 = 0.0283, wR_2 = 0.0630
Largest diff. peak/hole / e Å^{-3}	1.72/-0.93
Flack parameter	-0.04(2)

N-Boc-*trans*-4-(iodo)-L-proline methyl ester, Boc-*trans*-iodoproline-OMe. **43**



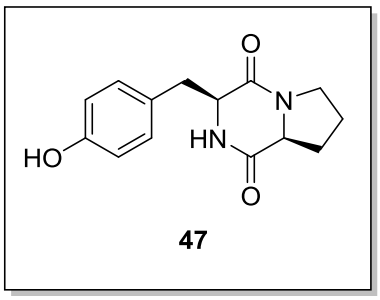
43

Boc-Hyp-OMe **38** (0.24 g, 0.96 mmol), imidazole (0.07 g, 1.06 mmol), PPh_3 (0.28 g, 1.06 mmol) and I_2 (0.27 g, 1.06 mmol) were dissolved in DCM (20 mL) and left to stir at room temperature for 48 h. The reaction mixture was evaporated under reduced pressure and purified *via* column chromatography (SiO_2; 100% hexane → 60/40%

hexane/EtOAc) yielding **43** as a sticky yellow solid (0.12 g, 35%). R_f 0.16 (hexane/EtOAc 9:1). δ_H mixture of two rotamers in a ratio of ca. 1.0:1.2 (400 MHz; $CDCl_3$) 1.42 and 1.46 (9H, 2 s, BocH), 2.41 (1H, m, β-CH), 2.59 (1H, m, β-CH), 3.68 to 3.74 (4H, m, OCH_3, δ-CH), 3.97 (1H, m, δ-CH), 4.32 to 4.47 (2H, m, γ-CH, α-CH); δ_C (100 MHz; $CDCl_3$) δ_C (100 MHz, $CDCl_3$) 172.9, 172.8, 153.9, 153.3, 80.8, 59.0, 58.6, 57.8, 57.5, 52.6, 52.4, 43.3, 42.3, 29.4, 28.5, 16.2, 15.9; *m/z* (ES^+) 378.0 $[M+Na]^+$.

Data obtained is consistent with that given in the literature: δ_H (250 MHz; $CDCl_3$) 1.42 and 1.47 (9H, 2 s), 2.41 (1H, m), 2.58 (1H, m), 3.72 and 3.80 (2H, 2 m), 3.74 (3H, s), 3.96 and 3.98 (1H, 2 m) 4.30 to 4.35 (2H, m).[12]

Cyclo(L-Pro-L-Tyr). **47**

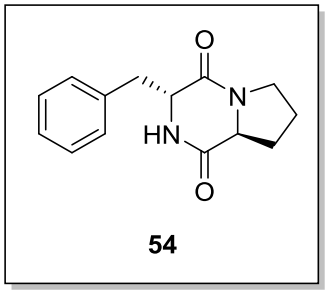


47

General procedure for cyclic dipeptide (DKP) synthesis was followed on a 0.30 mmol scale with amino acid (A): Boc-L-Pro-OH and (B): Boc-L-Tyr-OH to yield 47 as pale orange powder (31 mg, 46%). δ_H (400 MHz; $CDCl_3$) 1.86 (2H, m, Pro CH_2), 1.98 (1H, m, Pro-CH), 2.30 (1H, m, Pro-CH), 2.79 (1H, dd, J 9.4, 14.5, Tyr-ß-CH) 3.40 (1H, dd, J 4.0, 14.5, Tyr-ß-CH), 3.57 (2H, m, Pro-δ-CH_2), 4.06 (1H, m, Pro-α-CH), 4.23 (1H, m, Tyr-α-CH), 6.16 (1H, s, NH), 6.76 (2H, d, J 8.4, ArH), 7.02 (2H, d, J 8.4, ArH); δ_C (100 MHz; $CDCl_3$) 169.9, 165.4, 156.0, 130.5, 126.8, 116.2, 59.3, 56.4, 45.5, 36.1, 28.4, 22.5; HRMS m/z (ES^+) 261.1232 ($[M+H]^+$. $C_{14}H_{17}N_2O_3$ requires 261.1239).

Data obtained is consistent with that given in the literature: δ_H (400 MHz; $CDCl_3$) 1.83 to 1.94 (2H, m), 1.96 to 2.02 (1H, m), 2.29 (1H, m), 2.75 (1H, dd), 3.39 to 3.44 (1H, m), 3.51 to 3.56 (1H, m). 3.59 to 3.66 (1H, m), 4.08 (1H, t), 4.21 (1H, dd), 5.72 (1H, s), 6.77 (2H, d), 7.05 (2H, d).[13]

Cyclo(L-Pro-D-Phe). **54**

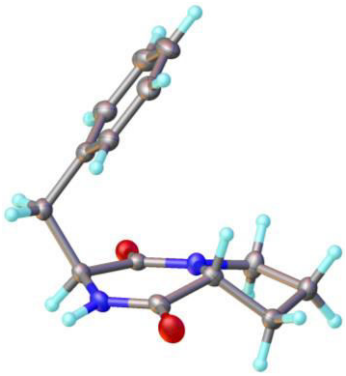


54

General procedure for cyclic dipeptide (DKP) synthesis was followed on a 0.05 mmol scale with amino acid (***A***): Boc-L-Pro-OH and (***B***): Boc-D-Phe-OH to yield cyclo(L-Pro-D-Phe) **54** as a pale yellow powder (6 mg, 49%). δ_H (400 MHz; $CDCl_3$) 1.74 (2H, m, Pro-γ-CH_2), 1.94 (1H, Pro-ß-CH), 2.20 (1H, Pro-ß-CH), 3.02 (1H, m, Pro-α-CH), 3.13 (2H, m, Phe-ß-CH_2), 3.40 (1H, m, Pro-δ-CH), 3.63 (1H, m, Pro-δ-CH), 4.23 (1H, m, Phe-α-CH), 6.26 (1H, bs, NH), 7.17 to 7.24 (2H, m, ArH), 7.27 to 7.35 (3H, m, ArH); δ_C (100 MHz; $CDCl_3$) 169.2, 164.9, 135.4, 130.0, 128.9, 127.7, 59.3, 57.9, 45.3, 40.7, 29.1, 21.4; HRMS *m/z* (ES^+) 245.1294 ($[M+H]^+$. $C_{14}H_{17}N_2O_2$ requires 245.1290).

Data obtained is consistent with that given in the literature: δ_H (400 MHz; d_4-methanol) 1.58 to 1.78 (2H, m), 1.84 to 2.00 (1H, m), 2.00 to 2.21 (1H, m), 2.65 (1H, dd, *J* 6.3, 10.4), 3.02 (1H, dd, *J* 4.6, 13.6), 3.23 (1H, dd, *J* 4.8, 13.7), 3.27 to 3.44 (1H, m), 3.49 to 3.65 (1H, m), 4.23 (1H, t, *J* 4.7), 7.14 to 7.44 (5H, m).[14]

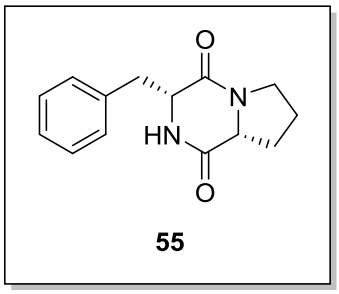
Crystal data and structure refinement:



Empirical formula	$C_{14}H_{16}N_2O_2$
Formula weight	244.29
Temperature/K	120
Crystal system	monoclinic
Space group	$P2_1$
a/Å	7.7670(5)
b/Å	6.5187(5)
c/Å	12.1777(9)
α/°	90
β/°	94.773(17)
γ/°	90
Volume/$Å^3$	614.43(8)
Z	2

ρ_{calc}g/cm^3	1.320
μ/mm^{-1}	0.090
F(000)	260.0
Crystal size/mm^3	0.48 × 0.17 × 0.05
Radiation	MoKα (λ = 0.71073)
2Θ range for data collection/°	3.356 to 60.02
Index ranges	$-10 \leq h \leq 10$, $-9 \leq k \leq 9$, $-17 \leq l \leq 17$
Reflections collected	8753
Independent reflections	3585 [R_{int} = 0.0315, R_{sigma} = 0.0363]
Data/restraints/parameters	3585/1/167
Goodness-of-fit on F^2	1.024
Final R indexes [I>=2σ (I)]	R_1 = 0.0383, wR_2 = 0.0983
Final R indexes [all data]	R_1 = 0.0437, wR_2 = 0.1014
Largest diff. peak/hole / e Å^{-3}	0.32/-0.18
Flack parameter	0.4(6)

Cyclo(D-Pro-D-Phe). **55**

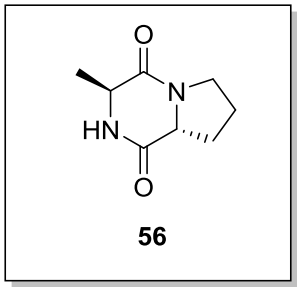


55

General procedure for cyclic dipeptide (DKP) synthesis was followed on a 0.08 mmol scale with amino acid (***A***): Boc-L-Pro-OH and (***B***): Boc-D-Phe-OH to yield cyclo(L-Pro-D-Phe) **55** as orange oil (7 mg, 42%). δ_H (400 MHz; $CDCl_3$) 1.84 to 2.08 (3H, m, Pro-CH_2, Pro-CH), 2.35 (1H, m, Pro-CH), 2.77 (1H, dd, *J* 10.8, 14.7, Phe-β-CH) (2H, dd, Pro-CH_2), 3.51 to 3.70, (3H, m, Pro-δ-CH_2, Phe-β-CH), 4.08 (1H, t, *J* 7.4, Pro-α-CH), 4.27 (1H, dd, *J* 3.6, 10.8, Phe-α-CH) 5.62 (1H, s, NH), 7.20 to 7.41 (5H, m, ArH); δ_C (100 MHz; $CDCl_3$) 169.5, 165.2, 136.1, 129.4, 129.3, 127.7, 59.3, 56.3, 45.6, 36.9, 28.5, 22.7; HRMS *m/z* (ES^+) 245.1310 ($[M+H]^+$. $C_{14}H_{17}N_2O_2$ requires 245.1290).

Data obtained is consistent with that given in the literature: δ_H (300 MHz; $CDCl_3$) 1.81 to 1.98 (2H, m), 2.00 to 2.10 (1H, m), 2.28 to 2.38 (1H, m), 2.77 (1H, dd), 3.45 to 3.59 (1H, m), 3.50 to 3.61 (1H, m), 3.61 to 3.69 (1H, m), 4.08 (1H, t), 4.27 (1H, dd), 5.62 (1H, br s), 7.18 to 7.41 (5H, m).[15]

Cyclo(D-Pro-L-Ala). **56**

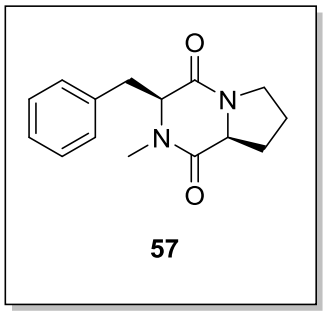


56

General procedure for cyclic dipeptide (DKP) synthesis was followed on a 0.16 mmol scale with amino acid (***A***): Boc-D-Pro-OH and (***B***): Boc-L-Ala-OH to yield cyclo(D-Pro-L-Ala) **56** as white solid (13 mg, 51%). δ_H (400 MHz; $CDCl_3$) 1.49 (3H, d, *J* 6.8, Ala-CH_3), 1.84 to 2.08 (3H, m, Pro-CH, Pro-CH_2), 2.41 (1H, m, Pro-CH), 3.53 (1H, m, CH), 3.64 (1H, m, CH), 3.99 to 4.12 (2H, m, Pro-α-CH, Ala-α-CH) 6.62 (1H, s, NH); δ_C (100 MHz; $CDCl_3$) 169.2, 166.5, 58.1, 53.9, 45.8, 29.2, 22.3, 20.4; HRMS *m/z* (ES^+) 169.0974 ($[M+H]^+$. $C_8H_{13}N_2O_2$ requires 169.0977).

Data obtained is consistent with that given in the literature: δ_H (400 MHz; $CDCl_3$) 1.48 (3H, d, *J* 7.2), 1.86 to 2.08 (3H, m), 2.41 (1H, m), 3.54 (1H, m), 3.66 (1H, m) 4.02 to 4.13 (2H, m), 7.40 (1H, br s).[16]

Cyclo(L-Pro-L-MePhe). **57**

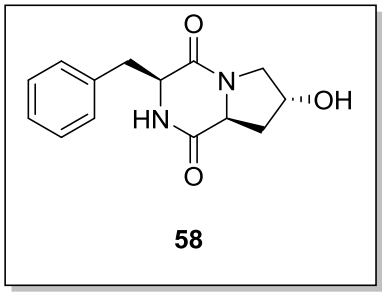


57

General procedure for cyclic dipeptide (DKP) synthesis was followed. However, microwave heating (60 °C, 10 min) was used in the final synthetic step (cyclisation). The synthesis was undertaken on a 0.10 mmol scale with amino acid (***A***): Boc-L-Pro-OH and (***B***): Boc-L-MePhe-OH to yield cyclo(L-Pro-L-MePhe) **57** as a pale orange powder (11 mg, 42%). δ_H (400 MHz; $CDCl_3$) 1.48 to 1.65 (3H, m, Pro-CH_2, Pro-CH), 1.86 (1H, m, Pro-CH), 3.10 to 3.20 (4H, m, NCH_3, CH), 3.40 (1H, m, CH), 3.70 (2H, m, CH_2) 4.28 (1H, m, CH), 7.03 to 7.08 (2H, m, ArH). 7.19 to 7.29 (3H, m, ArH); δ_C (100 MHz; $CDCl_3$) 169.2, 164.9, 135.4, 130.0, 129.0, 127.7, 77.4, 57.9, 45.3, 40.7, 29.3, 29.1, 21.9; HRMS *m/z* (ES^+) 259.1452 ($[M+H]^+$. $C_{15}H_{19}N_2O_2$ requires 259.1447).

Data obtained is consistent with that given in the literature: *m/z* (MALDI-TOF^+) 258.5 $[M+H]^+$.[17]

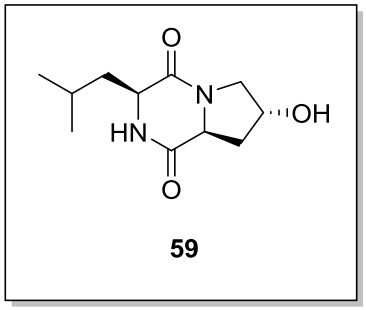
Cyclo(L-Hyp-L-Phe). **58**



58

General procedure for cyclic dipeptide (DKP) synthesis was followed on a 0.28 mmol scale with amino acid (***A***): Boc-L-Hyp-OH and (***B***): Boc-L-Phe-OH to yield cyclo(L-Hyp-L-Phe) **58** as white powder (62 mg, 86%). ν_{max}(solid)/cm^{-1} 3275, 2928 (w, C-H), 2871, 1637 (s, C=O), 1414; $[\alpha]^{D}_{25}$ = +104.0° (c 0.1, $CHCl_3$); δ_H (400 MHz; $CDCl_3$) 1.95 (1H, m, Hyp-CH), 2.32 (1H, dd, *J* 6.2, 13.5, Hyp-CH), 2.79 (1H, dd, *J* 10.4, 14.4, Phe-ß-CH), 3.53 to 3.62 (2H, m, Phe-β-CH, Hyp-CH), 3.73 (1H, dd, *J* 4.8, 13.2, Hyp-CH), 4.31 (1H, m, Phe-α-CH), 4.44 (1H, dd, *J* 6.0, 11.2, Hyp-CH), 4.52 (1H, t, *J* 4.2, Hyp-CH), 5.84 (1H, s, NH), 7.20 to 7.24 (2H, m, ArH), 7.27 to 7.37 (3H, m, ArH); δ_C (100 MHz; $CDCl_3$) 169.8, 165.3, 135.9, 129.4, 129.3, 127.7, 57.5, 56.3, 54.6, 37.8, 36.8; HRMS *m/z* (ES$^+$) 261.1219 ([M+H]$^+$. $C_{14}H_{17}N_2O_3$ requires 261.1239).

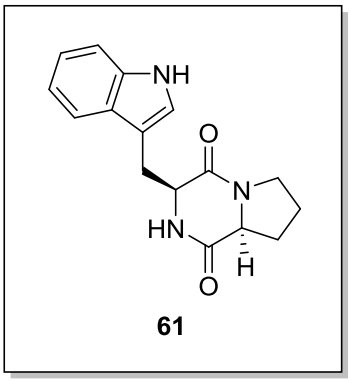
Cyclo(L-Hyp-L-Leu). **59**



59

General procedure for cyclic dipeptide (DKP) synthesis was followed on a 0.07 mmol scale with amino acid (***A***): Boc-L-Hyp-OH and (***B***): Boc-L-Leu-OH to yield cyclo(L-Hyp-L-Leu) **59** as white powder (5 mg, 31%). ν_{max}(solid)/cm^{-1} 3307 (br, O-H), 2963 (w, C-H), 1657 (s, C=O), 1435; $[\alpha]^{D}_{25}$ = -79.1° (c 0.2, $CHCl_3$); δ_H (400 MHz; MeOD) 0.95 (3H, m, Leu-δCH_3), 0.97 (3H, m, Leu-δCH_3), 1.51 (1H, m, CH), 1.90 (2H, m, CH), 2.09 (1H, m, CH), 2.28 (1H, m, CH) 3.43 (1H, d, *J* 12.8, CH), 3.66 (1H, dd, *J* 4.4, 12.8, CH), 4.18 (1H, m, CH), 4.46 (1H, t, *J* 4.4, CH), 4.52 (1H, m, CH); δ_C (100 MHz; $CDCl_3$) 170.7, 166.6, 69.0, 57.8, 54.8, 53.0, 39.0, 38.0, 25.2, 23.8, 21.7; HRMS *m/z* (ES$^+$) 227.1390 ([M+H]$^+$. $C_{11}H_{19}N_2O_3$ requires 227.1390).

Cyclo(L-Pro-L-Trp). **61**

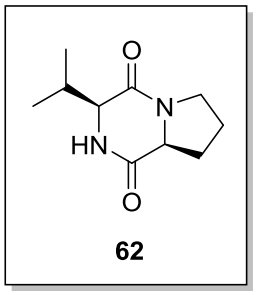


61

General procedure for cyclic dipeptide (DKP) synthesis was followed on a 0.27 mmol scale with amino acid (***A***): Boc-L-Pro-OH and (***B***): Boc-L-Trp-OH to yield cyclo(L-Pro-L-Trp) **61** as a white powder (51 mg, 67%). δ_H (400 MHz; $CDCl_3$) 1.80 to 2.07 (3H, m, Pro-CH_2, Pro-CH), 2.32 (1H, m, Pro-CH), 2.97 (1H, m, Trp-CH), 3.54 to 3.79 (3H, m, Pro-δ-CH_2, Trp-CH), 4.08 (1H, m, Pro-α-CH), 4.38 (1H, m, Trp-α-CH), 5.76 (1H, bs, amide-NH), 7.10 (1H, m, ArH), 7.14 (1H, m, ArH), 7.24 (1H, m, ArH), 7.39 (1H, d, *J* 8.0, ArH), 7.59 (1H, d, *J* 8.0, ArH), 8.33 (1H, br s, indole-NH); δ_C (100 MHz; $CDCl_3$) 169.5, 165.6, 136.8, 126.8, 123.5, 122.9, 120.1, 118.6, 111.7, 110.1, 59.4, 54.7, 45.6, 28.5, 27.0, 22.8; HRMS *m/z* (ES^+) 284.1400 ($[M+H]^+$. $C_{16}H_{18}N_3O_2$ requires 284.1399).

Data obtained is consistent with that given in the literature: δ_H (400 MHz; $CDCl_3$) 1.86 to 2.04 (3H, m), 2.27 to 2.35 (1H, m), 2.98 (1H, dd, *J* 10.8, 15.1), 3.55 to 3.69 (2H, m), 3.75 (1H, ddd, *J* 0.8, 3.6, 14.8), 4.05 (1H, t, *J* 7.6), 4.37 (1H, dd, *J* 2.4, 10.8), 5.78 (1H, br s), 7.08 (1H, d, *J* 2.1), 7.14 (1H, td, *J* 0.9, 7.5), 7.23 (1H, td, *J* 1.0, 7.5), 7.38 (1H, d, *J* 8.4), 7.59 (1H, d, *J* 8.0), 8.40 (1H, br s).[18]

Cyclo(L-Pro-L-Val). **62**



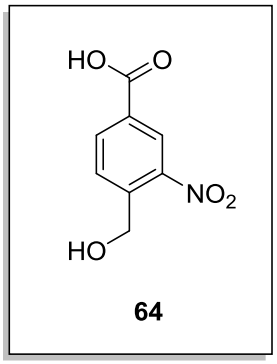
62

General procedure for cyclic dipeptide (DKP) synthesis was followed on a 0.15 mmol scale with amino acid (***A***): Boc-L-Pro-OH and (***B***): Boc-L-Val-OH to yield cyclo(L-Pro-L-Val) **62** as a white as white crystalline powder (12 mg, 41%). δ_H (400 MHz; $CDCl_3$) 0.91 (3H, d, *J* 6.8 ,CH_3), 1.06 (3H, d, *J* 7.2, CH_3), 1.90 (1H, m, CH), 2.04 (2H, m, CH), 2.37 (1H, m, CH), 2.64 (1H, m, CH), 3.54 (1H, m, CH), 3.63 (1H, m, CH), 3.94 (1H, m, CH), 4.08 (1H, m, CH), 5.74 (1H, m, NH); δ_C (100 MHz; $CDCl_3$) 170.1, 165.0, 60.5, 59.0, 45.3, 29.7, 29.5, 19.5, 16.2; HRMS *m/z* (ES^+) 197.1299 ($[M+H]^+$. $C_{10}H_{17}N_2O_2$ requires 197.1290).

Data obtained is consistent with that given in the literature: δ_H (500 MHz; $CDCL_3$) 0.92 (3H, d, *J* 7.6), 1.07 (3H, d, *J* 7.3), 1.87 to 1.95 (1H, m), 2.00 to 2.11 (1H, m), 2.35 to

2.41 (2H, m), 2.59 to 2.67 (1H, m), 3.51 to 3.68 (2H, m), 3.94 (1H, s), 4.08 (1H, t, *J* 7.6), 5.99 (1H, br s).[19]

4-hydroxymethyl-3-nitrobenzoic acid. **64**

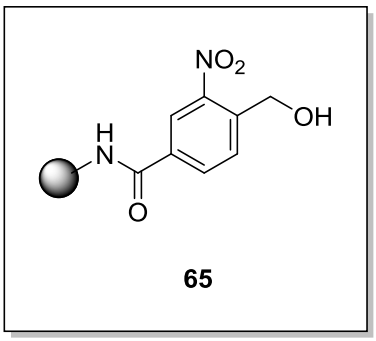


64

4-Bromomethyl-3-nitrobenzoic acid (2.00 g, 7.57 mmol) was dissolved in aqueous $NaHCO_3$ (50%, 30 mL) and heated to reflux for 40 min. The yellow solution was hot filtered and acidified to pH 1 – 2. The aqueous solution was extracted with EtOAc (3 x 30 mL) and the combined organic extracts were washed with saturated aqueous NaCl (30 mL). The organic layer was dried with $MgSO_4$, filtered and evaporated under reduced pressure to yield **64** as a pale orange powder (0.854 g, 59%); δ_H (400 MHz; $CDCl_3$) 5.01 (2H, m, CH_2), 8.01 (1H, m, ArH), 8.31 (1H, m, ArH), 8.63 (1H, m, ArH); δ_C (100 MHz; $CDCl_3$) 195.6, 176.6, 172.2, 163.3, 160.3, 158.0, 154.8, 90.0; *m/z* (ESI^-) 196.4 [M – H^+].

Data obtained is consistent with that given in the literature: δ_H (400 MHz; d_4-methanol) 4.99 (2H, s), 7.97 (1H, d, *J* 8.1), 8.27 (1H, dd, *J* 1.5, 8.1), 8.57 (1H, d, *J* 1.7).[20]

MHBA-linker, **65**

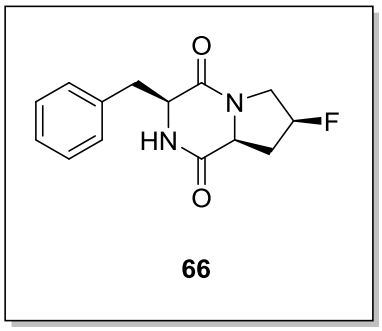


65

DCM (5 mL) was added to MHBA.LL.HCl resin (1.0 eq, 0.10 mmol) and the resin was agitated for 30 seconds before draining, this step was repeated a further 4 times. A solution of 40% TFA in DCM (5 mL) was added and the resin was agitated for 30 second before draining. Fresh 40% TFA in DCM (5 mL) was added and the resin was agitated for 5 min before draining, this step was repeated once more. A solution of 5% DIPEA in DCM (5 mL) was added and the resin was agitated for 2 min before draining, this step was repeated twice more. Finally, the resin was washed with DCM (5 x 5 mL). 4-hydroxymethyl-3-nitrobenzoic acid **01** (5.0 eq, 0.50 mmol), DIC (5.0 eq, 0.50 mmol) in DCM (1 mL) was added to MHBA.LL.HCl resin (1.0 eq, 0.10 mmol) and the solution was left to shake for 18 h. The resin was

washed with DMF (5 x 5mL) and DCM (5 x 5mL) to form pre-loaded MHBA-linker **65**.

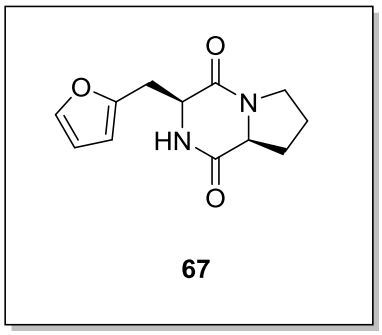
Cyclo(L-Flp-L-Phe). **66**



66

General procedure for cyclic dipeptide (DKP) synthesis was followed on a 0.10 mmol scale with amino acid (***A***): Boc-L-flp-OH **69** and (***B***): Boc-L-Phe-OH to yield cyclo(L-Flp-L-Phe) **66** as white powder (21 mg, 84%). ν_{max}(solid)/cm^{-1} 3168, 2944 (w, C-H), 2875, 1666 (s, C=O), 1418; $[\alpha]^{D}_{25}$ = -195.2° (c 0.5, $CHCl_3$); δ_H (400 MHz; $CDCl_3$) 2.50 (1H, m, flp-CH), 2.76 to 2.91 (2H, flp-CH, Phe-CH), 3.40 (1H, ddd, *J* 3.6, 13.6, 32.8, flp-CH), 3.65 (1H, dd, *J* 3.6, 14.6, Phe-CH), 4.20 (1H, dd, *J* 4.6, 10.2, flp-CH), 4.26 to 4.38 (2H, m, Phe-CH, flp-CH), 5.23 (1H, d, *J* 53.2, flp-γCH) 5.62 (1H, s, NH), 7.23 to 7.39 (5H, m, ArH); δ_C (100 MHz; $CDCl_3$) 168.6, 165.9, 135.8, 129.5, 129.2, 127.8, 89.9 (1C, d, *J* 179.5), 57.2, 56.1, 52.5 (1C, d, *J* 23.2), 36.5, 34.7 (1C, d, *J* 24.4); HRMS *m/z* (ES^+) 263.1770 ($[M+H]^+$. $C_{14}H_{16}FN_2O_2$ requires 263.1196).

Cyclo(L-Pro-L-β-(2-furyl)-Ala). **67**



67

To pre-loaded MHBA-linker **65** (1.0 eq, 0.10 mmol) a pre-mixed peptide coupling solution of Boc-L-Pro-OH (3.0 eq, 0.30 mmol), DMAP (0.05 eq) and DIC (3.0 eq, 0.30 mmol) in DCM (2 mL) was added. The reaction mixture was shaken for 40 min at room temperature. The solution was drained and coupling procedure was repeated. The resin was washed with DCM (5 x 5 mL). A deprotection solution of 40% TFA in DCM (3 mL) was added to the resin and stirred for 10 min. The solution was drained and this deprotection procedure was repeated twice. The resin was washed with DCM (5 x 5 mL) and stored at 5°C for 4 h. Fmoc-β-(2-furyl)-Ala-O*t*Bu **14** (4.0eq, 0.40 mmol) was dissolved in 20% TFA in DCM (10 mL) and stirred at room temperature for 3 hours or until the deprotection reaction was judged to be complete by TLC to yield Fmoc-β-(2-furyl)-Ala-OH (quant.). Fmoc-β-(2-furyl)-

Ala-OH was used without purification. A pre-mixed peptide coupling solution of Fmoc-β-(2-furyl)-Ala-OH (2.0 eq, 0.20 mmol), DIPEA (5.0 eq, 0.50 mmol) and PyBOP (2.0 eq, 0.20 mmol) in DCM (2 mL) was added to the resin and stirred for 1 h. The solution was drained and coupling procedure was repeated. The resin was washed with DCM (5 x 5 mL) and DMF (3 x 5 mL). Finally 20% piperidine in DMF (3 mL) was added to the resin and stirred for 10 min. The filtrate solution was drained, collected and then repeated. The combined filtrate solutions were evaporated under reduced pressure purified *via* column chromatography (SiO_2; 70/30% EtOAc/hexane → 90%/10% EtOAc/acetone) to yield **67** as colourless oil (14 mg, 64%). ν_{max}(solid)/cm^{-1} 3223, 2950 (w, C-H), 2879, 1646 (s, C=O), 1422; $[\alpha]^{D}_{23}$ = -259.5° (c 0.4, $CHCl_3$); δ_H (400 MHz; $CDCl_3$) 1.86 to 2.12 (3H, m, Pro-CH, Pro-CH_2) 2.35 (1H, m, Pro-CH), 2.90 (1H, dd, *J* 10.4, 15.6, FurylAla-β-CH), 3.52 to 3.66 (3H, m, FurylAla-β-CH, Pro-δ-CH_2) 4.10 (1H, m, Pro-α-CH), 4.27 (1H, d, *J* 10.4, FurylAla-α-CH), 6.06 (1H, s, NH), 6.19 (1H, dd, *J* 0.4, 3.2, ArH), 6.32 (1H, dd, *J* 2.0, 3.2, ArH), 7.36 (1H, dd, *J* 0.4, 2.0, ArH); δ_C (100 MHz, $CDCl_3$) 177.4, 169.6, 142.7, 150.6, 110.7, 108.6, 59.3, 54.7, 45.7, 28.4, 22.8, 29.3; HRMS *m/z* (ES^+) 235.1065 ($[M+H]^+$. $C_{12}H_{15}N_2O_3$ requires 235.1083).

7.2 Solid Phase Peptide Syntheses

7.2.1 Materials and Methods

Electrospray ionization (ESI) mass spectra were collected on a Waters TQD mass spectrometer and matrix assisted laser desorption/ionization - time of flight (MALDI-TOF) on a Waters Micromass LR TOF mass spectrometer or a Bruker Autoflex II ToF/ToF mass spectrometer (fitted with a 337 nm nitrogen lasers). Crude CCl2 peptide fragments were purified using a Gilson 305 LC pump and UV/Vis detector fitted with a C18 column: 5.0 μm, 250 mm x 10 mm. Absorbance was measured at 230 nm. The linear gradient ranged from 2% MeCN (0.1% TFA) in H_2O (0.1% TFA) at 0 min to 50% MeCN (0.1% TFA) in H_2O (0.1% TFA) at 60 min. All amino acid derivatives and peptide resins were purchased from Novabiochem. All chemicals were purchased from Sigma-Aldrich, unless otherwise stated. DMF and PyBOP were purchased from AGTC and Apollo Scientific, respectively. The amino acid derivatives that were utilised are as follows: Fmoc-Arg(Pbf)-OH, Fmoc-Asn(*t*Bu)-OH, Fmoc-Asp(*t*Bu)-OH, Fmoc-Cys(Trt)-OH, Fmoc-Gln(Trt)-OH, Fmoc-Glu(*t*Bu)-OH, Fmoc-His(Trt)-OH, Fmoc-

Lys(Boc)-OH, Fmoc-Ser(*t*Bu)-OH, Fmoc-Thr(*t*Bu)-OH, Fmoc-Trp(Boc)-OH and Fmoc-Tyr(*t*Bu)-OH.

7.2.2 General SPPS Procedures

7.2.2.1 Microwave Assisted Automated Fmoc-SPPS

All automated syntheses were undertaken on a CEM liberty1 microwave peptide synthesizer fitted with a Discover microwave unit in a 30 mL PTFE vessel. Pre-synthesis swelling of the resin was carried out in DMF for 1 h at room temperature. Microwave assisted coupling cycles were carried out for 10 min at 75°C (25 W) with amino acid (5.0 eq.), DIC (5.0 eq.) and HOBt (10.0 eq.). In the case of double couplings, the vessel was drained after the first cycle and the coupling cycle was repeated with fresh reagents. For amino acid and sequence specific details see the relevant section in **Chapter 5.** The Fmoc deprotection step was carried out for 10 min at 75°C (45 W) with a solution of 20% piperidine in DMF. In all steps, agitation was provided by bubbling nitrogen.

7.2.2.2 Room Temperature Manual Fmoc-SPPS

Manual couplings were carried out at room temperature for 1h 30 minutes with amino acid (5.0 eq.), PyBOP (5.0 eq.) and DIPEA (10.0 eq.) in a fritted polypropene vessel. In the case of an incomplete coupling (visualised by TNBS/Kaiser or chloranil tests), double or triple couplings were undertaken. For successive couplings the vessel was drained after the each coupling cycle and the coupling cycle was repeated with fresh reagents. For amino acid and sequence specific details see the relevant section in **Chapter 5.** The Fmoc deprotection step was carried out for 5 min with a solution of 20% piperidine in DMF, followed by a further treatment for 15 min with a fresh solution of 20% piperidine in DMF. In all steps, agitation was provided by rolling.

7.2.2.3 *N*-Alloc Protection and Dawson Dbz Resin Loading

The Dawson Dbz (3-(Fmoc-amino)-4-aminobenzoyl) AM resin was swollen in DMF for 1 h at room temperature. The resin was treated with a solution of 20% piperidine in DMF and washed with DMF. The general procedure for room temperature manual Fmoc-SPPS was followed to couple to first amino acid to the resin. The peptide-resin was washed with DMF before treatment with a solution of allyl chloroformate (10.0 eq.) and DIPEA (1.0 eq.) in DCM (minimum volume to cover resin) for 24 h at room temperature. The peptide-resin was washed with DCM and DMF.

7.2.2.4 *N*-Alloc Deprotection

Pre-deprotection, DCM was added to the peptide-resin in a fritted polypropene vessel and the mixture was left to swell for 30 min followed by argon sparging for 30 sec. $Pd(PPH_3)_4$ (0.4 eq.) and $PhSiH_3$ (20.0 eq) were added in DCM (minimum volume to cover resin) and the reaction was rolled at room temperature for 30 min in an argon atmosphere. Subsequently, the peptide-resin was washed with DCM then DMF.

7.2.2.5 Nbz Formation

The peptide-resin in a fritted polypropene vessel was treated with *p*-nitrophenyl chloroformate (5.0 eq.) in DCM (minimum volume to cover resin) for 1 h at room temperature. Subsequently, the peptide-resin was washed with DCM, DMF and rolled in a solution of DIPEA (0.5 M) in DMF for 45 min at room temperature. Finally, the resin was washed with DMF.

7.2.2.6 Peptide Cleavage from Acid-Labile Resin

Pre-cleavage the peptide-resins are washed with DCM (x 3) and ether (x 3) and left to air dry for 5 min. Subsequently the peptide-resins were treated with a (4 mL per 0.1 mmol of resin) solution of 2.5% TIPS and 2.5% H_2O in TFA for 3 h at room temperature. After which, the resin was filtered and the filtrate collected. The filtrate was treated with cold ether to form a suspension that is left to settle or pelleted *via*

centrifugation. The precipitate was collected and washed with ether (x 2) yielding crude product peptide. For sequence specific cleavage details see the relevant section in **Chapter 5**.

7.2.2.7 Reductive Peptide Cleavage from Acid-Labile Resin

Pre-cleavage the peptide-resins are washed with DCM (x 3) and ether (x 3) and left to air dry for 5 min. Subsequently the peptide-resins were treated with a (4 mL per 0.1 mmol of resin) solution of 5% TIPS in TFA for 2 h 30 min at room temperature. After which, TMSBr (60 μL per 0.1 mmol of resin) and EDT (70 μL per 0.1 mmol of resin) were added and the solution was left for 30 min at room temperature. The resin was filtered and the filtrate collected. The filtrate was treated with cold ether to form a suspension that is left to settle or pelleted *via* centrifugation. The precipitate was collected and washed with ether (x 2) yielding crude product peptide. For sequence specific cleavage details see the relevant section in **Chapter 5**.

7.2.3 Syntheses of CCL2 Peptide Fragments

CCL2 Peptide Fragment **52-76**

General procedure for microwave assisted automated Fmoc-SPPS synthesis was followed on a 0.1 mmol scale with low-loading polystyrene Wang (PABA) resin (0.27 mmol/g). Fmoc-His(Trt)-OH coupling was carried out at a reduced temperature (10 min at 50 °C (25 W). The general procedure for peptide cleavage from acid-labile resin was followed and crude CCL2 Peptide Fragment **52-76** was isolated at a purity of 80.2% by RP-HPLC (230 nm) subsequent purification *via* RP-HPLC afforded CCL2 Peptide Fragment **52-76** (50 mg, 17%). *m/z* (MALDI-TOF$^+$) 2928.8 ([M+H]$^+$. $C_{124}H_{199}N_{36}O_{42}S_2$ requires 2928.4). Retention time by analytical RP-HPLC (230 nm) of 16.2 minutes (gradient: 2% MeCN in H_2O (0.1% TFA) to 50% MeCN in H_2O (0.1% TFA) over 30 minutes).

CCL2 Peptide Fragment **36-51**

General procedure for microwave assisted automated Fmoc-SPPS synthesis was followed on a 0.1 mmol scale with Dawson Dbz AM resin (0.41 mmol/g) and in every case the amino acid coupling cycle was repeated. General procedures for *N*-Alloc protection, *N*-Alloc deprotection and Nbz formation were followed**.** Boc-L-thiazolidone-4-carboxylic acid coupling was carried out following the general procedure for room temperature manual Fmoc-SPPS with reduced reagent equivalents: amino acid (2.5 eq.), PyBOP (2.5 eq.) and DIPEA (5.0 eq.) The general procedure for peptide cleavage from acid-labile resin was followed and crude CCL2 Peptide Fragment **36-51** was isolated at a purity of 32.7% by RP-HPLC (230 nm). Subsequent purification *via* RP-HPLC afforded CCL2 Peptide Fragment **36-51** (23 mg, 14%). *m/z* (MALDI-TOF$^+$) 1960.4 ([M+H]$^+$. $C_{92}H_{147}N_{22}O_{23}S$ requires 1960.1). Retention time by analytical RP-HPLC (230 nm) of 22.4 minutes (gradient: 2% MeCN in H_2O (0.1% TFA) to 50% MeCN in H_2O (0.1% TFA) over 30 minutes).

CCL2 Peptide Fragment **36-76**

General procedure for microwave assisted automated Fmoc-SPPS synthesis was followed on a 0.1 mmol scale with pre-loaded (Thr(*t*Bu)) Tentagel resin (0.21 mmol/g) (Rapp-Polymere) and in every case the amino acid coupling cycle was repeated. Higher temperature couplings (10 min at 95°C (25 W) were utilised for section 46-51 (residues: IVAKEI) and section 37-43 (residues: PKEAVI). Fmoc-Cys(Trt)-OH and Fmoc-His(Trt)-OH couplings were carried out at a reduced temperature (10 min at 50 °C (25 W). Fmoc-Lys(Boc)-Thr(Ψpro)-OH (KT pseudo-proline) coupling was carried out following the general procedure for room temperature manual Fmoc-SPPS with reduced reagent equivalents: amino acid (2.5 eq.), PyBOP (2.5 eq.) and DIPEA (5.0 eq.) and repeated until complete (as visualised by TNBS test). The general procedure for reductive peptide cleavage from acid-labile resin was followed and crude CCL2 Peptide Fragment **36-76** was isolated at a purity of 33.1% by RP-HPLC (230 nm). Subsequent purification *via* RP-HPLC afforded CCL2 Peptide Fragment **36-76** (6 mg, 8%). *m/z* (ES$^+$) 941.1 ([M+5H]$^{5+}$. $C_{207}H_{342}N_{55}O_{63}S_3$ requires 940.5). Retention time by analytical RP-HPLC (230 nm) of 21.0 minutes (gradient: 2% MeCN in H_2O (0.1% TFA) to 50% MeCN in H_2O (0.1% TFA) over 30 minutes).

CCL2 Peptide Fragment **1-35**

General procedure for microwave assisted automated Fmoc-SPPS synthesis was followed on a 0.1 mmol scale with Dawson Dbz AM resin (0.41 mmol/g) and in every case the amino acid coupling cycle was repeated. Fmoc-Cys(Trt)-OH coupling was carried out at a reduced temperature (10 min at 50 °C (25 W). Fmoc-Arg(Pbf)-OH coupling was carried out at a reduced temperature (1 h min at room temperature). General procedures for *N*-Alloc protection, *N*-Alloc deprotection and Nbz formation were followed. Boc-L-pyroglutamic acid coupling was carried out following the general procedure for room temperature manual Fmoc-SPPS with reduced reagent equivalents: amino acid (2.5 eq.), PyBOP (2.5 eq.) and DIPEA (5.0 eq.) and repeated until complete (as visualised by a chloranil test). The general procedure for peptide cleavage from acid-labile resin was followed and crude CCL2 Peptide Fragment **1-35** was isolated at a purity of 19.8% by RP-HPLC (230 nm). Subsequent purification *via* RP-HPLC afforded CCL2 Peptide Fragment **1-35** (6 mg, 1%). *m/z* (ES^+) 1381.2 ($[M+3H]^{3+}$. $C_{180}H_{287}N_{56}O_{53}S_2$ requires 1381.7). Retention time by analytical RP-HPLC (230 nm) of 18.1 minutes (gradient: 2% MeCN in H_2O (0.1% TFA) to 50% MeCN in H_2O (0.1% TFA) over 30 minutes).

Nitrated (Tyr(NO_2)28) CCL2 Peptide Fragment **25-35**

General procedure for microwave assisted automated Fmoc-SPPS synthesis was followed on a 0.1 mmol scale with Dawson Dbz AM resin (0.41 mmol/g) and in every case the amino acid coupling cycle was repeated. Fmoc-Arg(Pbf)-OH coupling was carried out at a reduced temperature (1 h min at room temperature). General procedure for *N*-Alloc protection was followed. Fmoc-L-3-nitrotyrosine coupling was carried out following the general procedure for room temperature manual Fmoc-SPPS and repeated (visualisation by TNBS/Kaiser or chloranil proved test). The general procedure for peptide cleavage from acid-labile resin was followed and afforded nitrated (Tyr(NO_2)28) CCL2 peptide fragment **25-35.** *m/z* (MALDI-TOF^+) (1543.8 $[M+H]^+$. $C_{66}H_{107}N_{22}O_{21}$ requires 1543.8). Retention time by analytical RP-HPLC (230 nm) of 15.7 minutes (gradient: 2% MeCN in H_2O (0.1% TFA) to 50% MeCN in H_2O (0.1% TFA) over 30 minutes).

7.3 Chemotaxis Assays

7.3.1 Materials and Method

Cellular migration of THP-1 (derived from the peripheral blood of a 1 year old child with acute monocytic leukemia (ATCC TIB-202)) was measured in dual chamber 24-well companion plates (Becton Dickinson, Franklin Lakes, NJ) with 3μM filters. Prior to the experiment each companion plates was incubated at room temperature with a solution of 1% BSA/RPMI (1ml). The BSA solution was removed and to the lower well 800 μL of serum-free RPMI-1640 medium supplemented with CCL2 (10 nM). The DKP compound under investigation was also added to the lower cell (100 μM), except in the positive control. 500000 monocytes in a 500 μL solution of 1% bovine serum albumin (BSA) in RPMI 1640 were added to the upper well and the assay was incubated at 37 °C for 90 min before removal of medium. The upper surface of the transwell filters were gently rubbed to remove non-migrated cells and then fixed (1 h in cold methanol followed by 30 min staining *via* haematoxylin (Gill No.1; Sigma Aldrich)). The filters were dehydrated, mounted to slides and the migrant cells (adherent to the inferior aspect of the membrane) in five high-power fields of vision were counted *via* high resolution microscopy. The assays were carried out in triplicate.

7.4 References

1. Y. Koseki, H. Yamada and T. Usuki, *Tetrahedron: Asymmetry*, 2011, **22**, 580-586.

2. H. J. C. Deboves, C. A. G. N. Montalbetti and R. F. W. Jackson, *J. Chem. Soc., Perkin Transactions 1*, 2001, 1876-1884.
3. F. Friscourt, C. J. Fahrni and G.-J. Boons, *J. Am. Chem. Soc.*, 2012, **134**, 18809-18815.
4. H. Liu, V. R. Pattabiraman and J. C. Vederas, *Organic letters*, 2007, **9**, 4211-4214.
5. M. J. O'Donnell and F. Delgado, *Tetrahedron*, 2001, **57**, 6641-6650.
6. S. J. Cowling, K. J. Toyne and J. W. Goodby, *J. Mat. Chem.*, 2001, **11**, 1590-1599.
7. K. F. Stensrud, D. Heger, P. Šebej, J. Wirz and R. S. Givens, *Photochem. Photobiol. Sci.*, 2008, **7**, 614-624.
8. M. Yar, E. M. McGarrigle and V. K. Aggarwal, *Angew. Chem.-Int. Edit.*, 2008, **47**, 3784-3786.
9. A. Mollica, M. P. Paradisi, K. Varani, S. Spisani and G. Lucente, *Bioorg. Med. Chem.*, 2006, **14**, 2253-2265.
10. A. Banerjee and V. A. Kumar, *Bioorg. Med. Chem.*, 2013, **21**, 4092-4101.
11. K. K. Schumacher, J. Jiang and M. M. Joullié, *Tetrahedron: Asymmetry*, 1998, **9**, 47-53.
12. J. R. Dormoy, *Synthesis*, 1982, 753-756.
13. H. Thajudeen, K. Park, S.-S. Moon and I. S. Hong, *Tetrahedron Lett.*, 2010, **51**, 1303-1305.
14. H. Kries, R. Wachtel, A. Pabst, B. Wanner, D. Niquille and D. Hilvert, *Angew. Chem.-Int. Edit.*, 2014, **53**, 10105-10108.
15. F. Fdhila, V. Vazquez, J. L. Sanchez and R. Riguera, *J. Nat. Prod.*, 2003, **66**, 1299-1301.
16. J. Baek, S. Y. Kang, C. Im and Y. S. Park, *Eur. J. Org. Chem.*, 2014, **2014**, 2780-2789.
17. M. Teixido, E. Zurita, M. Malakoutikhah, T. Tarrago and E. Giralt, *J. Am. Chem. Soc.*, 2007, **129**, 11802-11813.
18. E. M. Boyd and J. Sperry, *Org. Lett.*, 2014, **16**, 5056-5059.
19. S. K. Nishanth, B. Nambisan and C. Dileep, *Peptides*, 2014, **53**, 59-69.
20. A. Eisenfuhr, P. S. Arora, G. Sengle, L. R. Takaoka, J. S. Nowick and M. Famulok, *Bioorg. Med. Chem.*, 2003, **11**, 235-249.

Appendices

Appendix 1: 4-Fluoroproline Synthesis

A1.1 The Influence of Fluoroproline

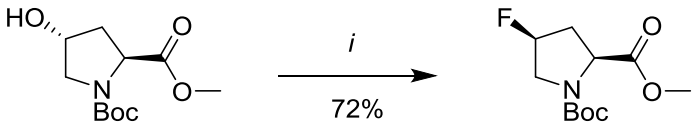
Cis(2*S*,4*S*)- and *trans*- 4-Fluoroproline (**Figure A1.1**) are of interest amino acids as they have been shown to drastically increase structural rigidity of peptides[1-3] and proteins.[4-6] The enhancement in structural rigidity arises from fluorine gauche and n→π* stereoelectronic effects.[7-10] The effects lead to a preference for the *cis* and *trans* diastereoisomers to form a C^{γ}-*endo* or C^{γ}-*exo* pyrrolidine ring pucker and in turn a *cis* and *trans* peptide bond, respectively. Hence, it was envisaged that these structural properties would have effects on the biological activity of both lipopeptides (**Section 2.4**) and diketopiperazine (DKP) CCL2 induced chemotaxis inhibitors (**Chapter 4**).

cis-fluoroproline trans-fluoroproline

Figure A1.1: *cis*(2*S*,4*S*)- and *trans*-(2*S*,4*R*)- fluoroproline.

A1.2 Synthesis of cis-Fluoroproline

Cis-fluoroproline (flp) and *trans*-fluoroproline (Flp) are expensive commercially available compounds (Fluorochem, Bachem). However, facile routes to their synthesis from the significantly cheaper *trans*-hydroxyproline (Hyp) (53$ per 100g, mimotopes) amino acid exist. Therefore, a published synthetic route to yield *cis*-fluoroproline was successfully followed from Boc-L-Hyp-OMe (**Scheme A1.1**). The synthesis of *trans*-fluoroproline is more difficult as it involves stereochemical inversion at the 4-position before final fluorination (as in **Scheme A1.1**). Therefore, the synthesis of *trans*-fluoroproline was not attempted.



38
Boc-Hyp-OMe

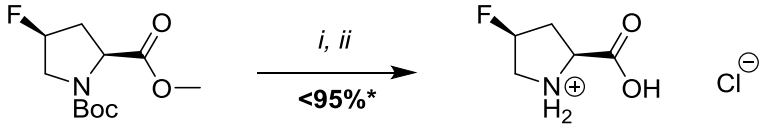
68
Boc-flp-OMe

Scheme A1.1: Fluorination of hydroxyproline *i*:TBAF, PBSF, DIPEA, DCM, RT, 48 h.

A1.3 Incorporation of cis-Fluoroproline into Bioactive Peptides

A1.3.1 Incorporation into Lipopeptides

The *cis*-fluoroproline (**68**) was deprotected to yield an unprotected amino acid hydrochloride salt (**30**) suitable for biosynthesis (**Scheme A1.2**).



68
Boc-flp-OMe

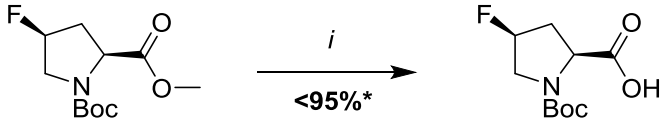
30
H-flp-OH.HCl

Scheme A1.2: Deprotection of **68**. *i*: LiOH, H_2O, THF, RT, 18 h. *ii*: HCl (aq), THF, RT, 5h.

The fully deprotected fluoroproline (**30**) was then supplemented into the growth medium of *Bascillus* sp. CS93 and extracts from the cultures were analysed for fluorinated lipopeptide content. However, no fluorinated lipopeptide products were observed

A1.3.2 Incorporation into DKP CCL2 Induced Chemotaxis Inhibitors

The on-resin synthetic route to form DKPs (**Chapter 4**) utilises *N*-Boc amino acids. Hence, Boc-flp-OH (**69**) was synthesised (**Scheme A1.3**).

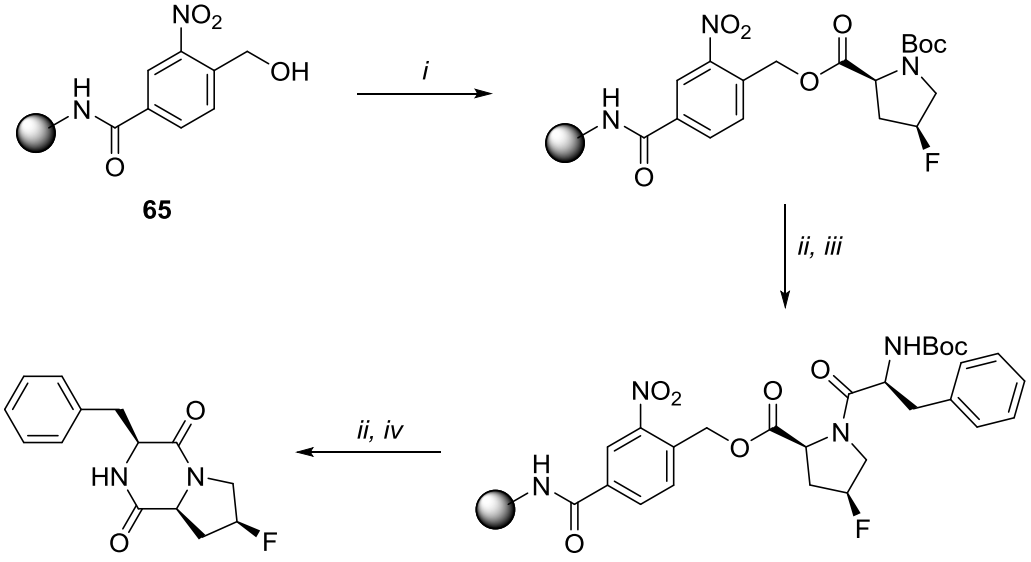


68
Boc-flp-OMe

69
Boc-flp-OH

Scheme A1.3: Methyl ester de-protection of *i*: LiOH, H_2O, THF, RT, 18 h

69 could easily be incorporated into a potentially active DKP: cyclo(flp-L-Phe) **66** following the general route described in **Chapter 4** and **Scheme A1.4**.



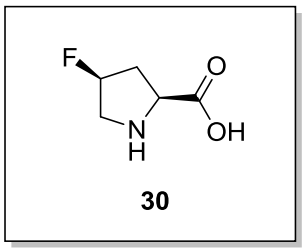
65

66
cyclo(flp-L-Phe)
84%*

Scheme A1.4: Synthesis of cyclo(flp-L-Phe) **66.** *i*: **69**, DIC, DMAP, DCM, 30 min, RT, x 2. *ii*: TFA:DCM (40% v/v), 10 min, RT, x 2. *iii*: Boc-Phe-OH, PyBOP, DIPEA, DCM, 1 h, RT, x 2. *iv*: DIPEA:DCM (10% v/v) 10 min, RT, x 3.*Denotes overall isolated yield over 5 steps.

A1.4 Experimental

H-flp-OH.HCl. **30**

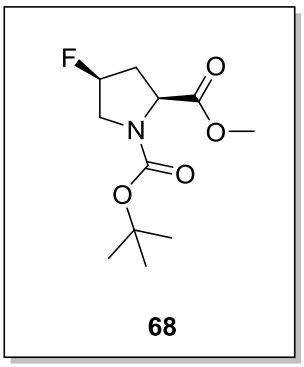


30

Boc-L-flp-O*t*Bu (0.09 g, 0.31 mmol) was dissolved in THF (5 mL) and stirred at room temperature. 0.5 mL of 11.6 M HCl (aq) was added dropwise and the reaction was left to stir for 30 min or until reaction completion was indicated by TLC. Solvent was evaporated under reduced pressure to yield **30** as a pale yellow powder (0.05 g, 90%); δ_H (400 MHz; d_4-methanol) 2.64 (2H, m, CH_2), 3.53 (1H, m, CH), 3.72 (1H, m, CH) 4.58 (1H, m, CH), 5.42 (1H, m, CH); δ_F (376 MHz; d_4-methanol) -175.8 (1F, m); δ_C (150 MHz; d_4-methanol) 170.9, 93.7, 91.9, 59.6, 53.5, 36.8; *m/z* (ES^-) 131.9 $[M–H]^-$.

Data obtained are consistent with that given in the literature: δ_H (400 MHz; $CDCl_3$) 2.54 to 2.38 (2H, m), 3.43 (1H, ddd, *J* 3.2, 13.8, 37.8), 3.72 (1H, dd, *J* 13.9, 18.3), 4.65 to 4.79 (1H, m), 5.34 (1H, d, *J* 52.2).[11]

N-Boc-*cis*-4-fluoro-L-proline methyl ester, Boc-flp-OMe. **68**

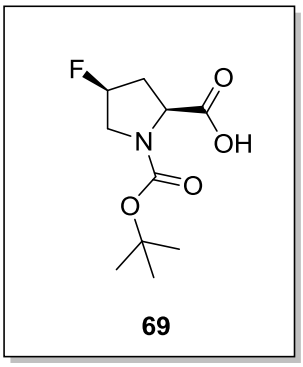


68

Boc-Hyp-OMe **38** (1.00 g, 4.09 mmol), DIPEA (1.76 mL, 10.22 mmol), TBAT (1.75 g, 3.27 mmol), PBSF (2.69 g, 9.99 mmol) in DCM (30 ml) was stirred for 48 h. The solvent was removed under reduced pressure and the crude product was purified *via* column chromotography (SiO_2; 100% hexane → 80/20% hexane/EtOAc) to yield **68** as a colourless oil (0.72 g, 72%), δ_H mixture of two rotamers in a ratio of ca. 1.0:1.4, (400 MHz; $CDCl_3$;) 1.43 and 1.48 (9H, 2 s, BocH), 2.39 (2H, m, β-CH_2), 3.54 to 3.92 (5H, m, δ-CH_2, OCH_3) 4.41 and 4.45 (1H, 2 d, *J* 9.6, α-CH), 5.19 (1H, d, *J* 52.8, γ-CH); δ_F (376 MHz; $CDCl_3$) -173.2 (1F, m, CF); δ_C (100 MHz; $CDCl_3$) 172.4, 172.0, 154.1, 153.8, 80.6, (91.3 and 92.3, 1C, 2 d, *J* 176, γ-C), 80.6, 57.8, 57.4, 52.5, 53.2 (1C, d, *J* 24, CH_2), 52.4, 37.2 (1C, d, *J* 22, CH_2), 28.5, 28.4; *m/z* (ES^+) 270.5 $[M+Na]^+$.

Data obtained is consistent with literature: δ_H a mixture of rotamers (400 MHz; $CDCL_3$) 1.44 and 1.49 (9H, s) 2.25 to 2.55 (2H, m), 3.56 to 3.93 (2H, m), 3.75 (3H, m), 4.43 and 4.55 (1H, d, *J* 9.2), 5.20 (br d, *J* 52.8).[12]

N-Boc-cis-4-fluoro-L-proline, Boc-flp-OH. **69**



69

Boc-flp-OMe **68** (0.80 g, 3.21 mmol) and LiOH (0.08 g, 3.2 mmol) in THF:H_2O (10 ml:5 ml) were stirred at room temperature for 48 h. The solvent was removed under reduced pressure and residue acidified to pH 4 with citric acid (10% w/v) and extracted with EtOAc (3 x 40 mL). The combined organic phases were dried over $MgSO_4$ and the solvent removed under reduced pressure to afford **69** as a clear oil (0.51 g, 68%); δ_H (400 MHz; $CDCl_3$) 1.40 (9H, s, *t*Bu) 2.13 to 2.48 (2H, m, β-CH_2), 3.49 to 3.82 (2H , m, δ-CH_2), 4.42 (1H, m, α-CH), 5.17 (1H, d, *J* 52.6, γ-CHF). δ_F (376 MHz; $CDCl_3$) -175.8 (1F, m, C-F); δ_C (100 MHz; $CDCl_3$) 174.8, 171.5, 73.7, 58.3, 43.2, 28.5.

Data obtained is consistent with literature: δ_H (400 MHz; d_4-methanol) 1.47 (9H, s), 2.39 to 2.46 (2H, m), 3.59 to 3.72 (2H, m), 4.38 to 4.43 (1H, m), 5.12 to 5.30 (1H, m).[13]

A1.5 References

1. D. D. Staas, K. L. Savage, V. L. Sherman, H. L. Shimp, T. A. Lyle, L. O. Tran, C. M. Wiscount, D. R. McMasters, P. E. Sanderson, P. D. Williams, B. J. Lucas, Jr., J. A. Krueger, S. D. Lewis, R. B. White, S. Yu, B. K. Wong, C. J. Kochansky, M. R. Anari, Y. Yan and J. P. Vacca, *Bioorg. Med. Chem.*, 2006, **14**, 6900-6916.
2. J. C. Horng and R. T. Raines, *Protein Sci.*, 2006, **15**, 74-83.
3. T. Steiner, P. Hess, J. H. Bae, B. Wiltschi, L. Moroder and N. Budisa, *PLoS One*, 2008, **3**, e1680.
4. W. Kim, R. A. McMillan, J. P. Snyder and V. P. Conticello, *J. Am. Chem. Soc.*, 2005, **127**, 18121-18132.
5. C. Renner, S. Alefelder, J. H. Bae, N. Budisa, R. Huber and L. Moroder, *Angew. Chem. -Int. Ed.*, 2001, **40**, 923-925.
6. D. Naduthambi and N. J. Zondlo, *J. Am. Chem. Soc.*, 2006, **128**, 12430-12431.
7. R. T. Raines, *Protein Sci.*, 2006, **15**, 1219-1225.
8. E. S. Eberhardt, N. Panisik, Jr. and R. T. Raines, *J. Am. Chem. Soc.*, 1996, **118**, 12261-12266.
9. M. P. Hinderaker and R. T. Raines, *Protein Sci.*, 2003, **12**, 1188-1194.
10. J. A. Hodges and R. T. Raines, *Org. Lett.*, 2006, **8**, 4695-4697.
11. R. Bejot, T. Fowler, L. Carroll, S. Boldon, J. E. Moore, J. Declerck and V. Gouverneur, *Angew. Chem.-Int. Edit.*, 2009, **48**, 586-589.
12. K. Y. Kim, B. C. Kim, H. B. Lee and H. Shin, *J. Org. Chem.*, 2008, **73**, 8106-8108.
13. X. Ji, M. Su, J. Wang, G. Deng, S. Deng, Z. Li, C. Tang, J. Li, J. Li, L. Zhao, H. Jiang and H. Liu, *Eur. J. Biochem.*, 2014, **75**, 111-122.

Appendix 2: Ciliatamide B Synthesis

A2.1 Anti-Parasitic Lipopeptides: The Ciliatamides

Ciliatamides A (**70**), B (**71**) and C (**72**) were isolated from deep sea sponge *Aaptos ciliate* by Nakao *et al.* as the (*S*,*S*)-enantiomers (**Figure A2.1**). These molecules are of particular interest as ciliatamide A and B show significant activity against insect-vector-borne *Leishmania* species.[1] Leishmaniasis is a neglected tropical disease that affects over 12 million people with an annual death toll exceeding 50,000.[2,3] The current treatment for leishmaniasis involves the use of drugs (pentavalent antimonials, miltefosene and amphotericin B) that possess a series of disadvantages including high cost, cardiotoxicity, parenteral administration and long treatment regimes.[4]

The solid phase approach described in **Chapter 4** is tolerant to a variety of chemical reactions, functional groups and can lead to the synthesis of complex DKP-natural product derivatives.[5] Hence, we envisaged the "on-resin" cyclisation synthesis of cyclic natural product lipopeptides, specifically: Ciliatamides B.[6]

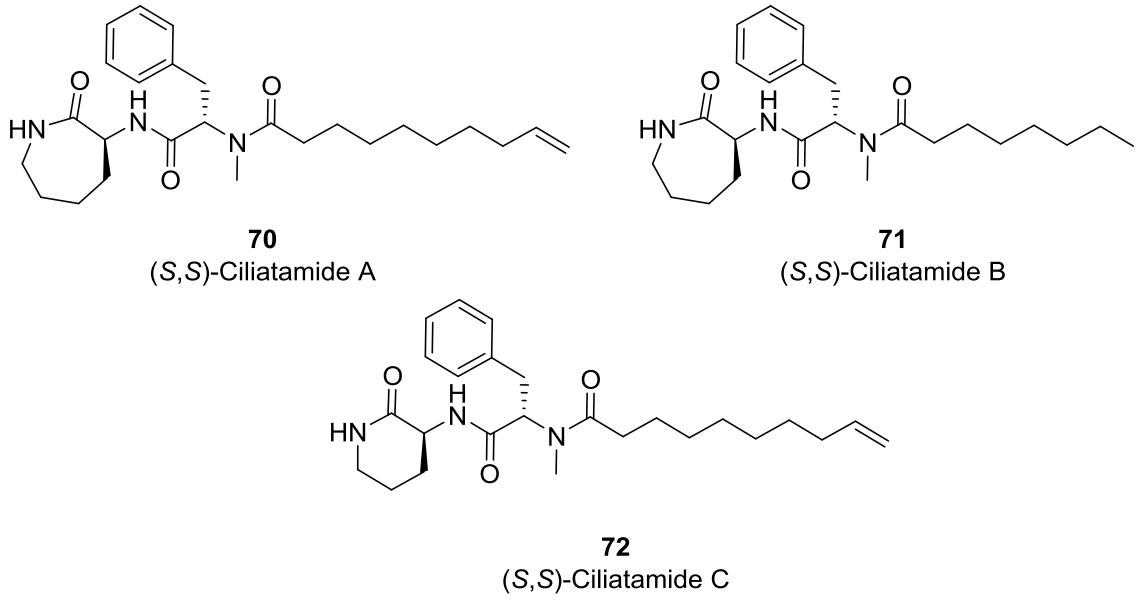
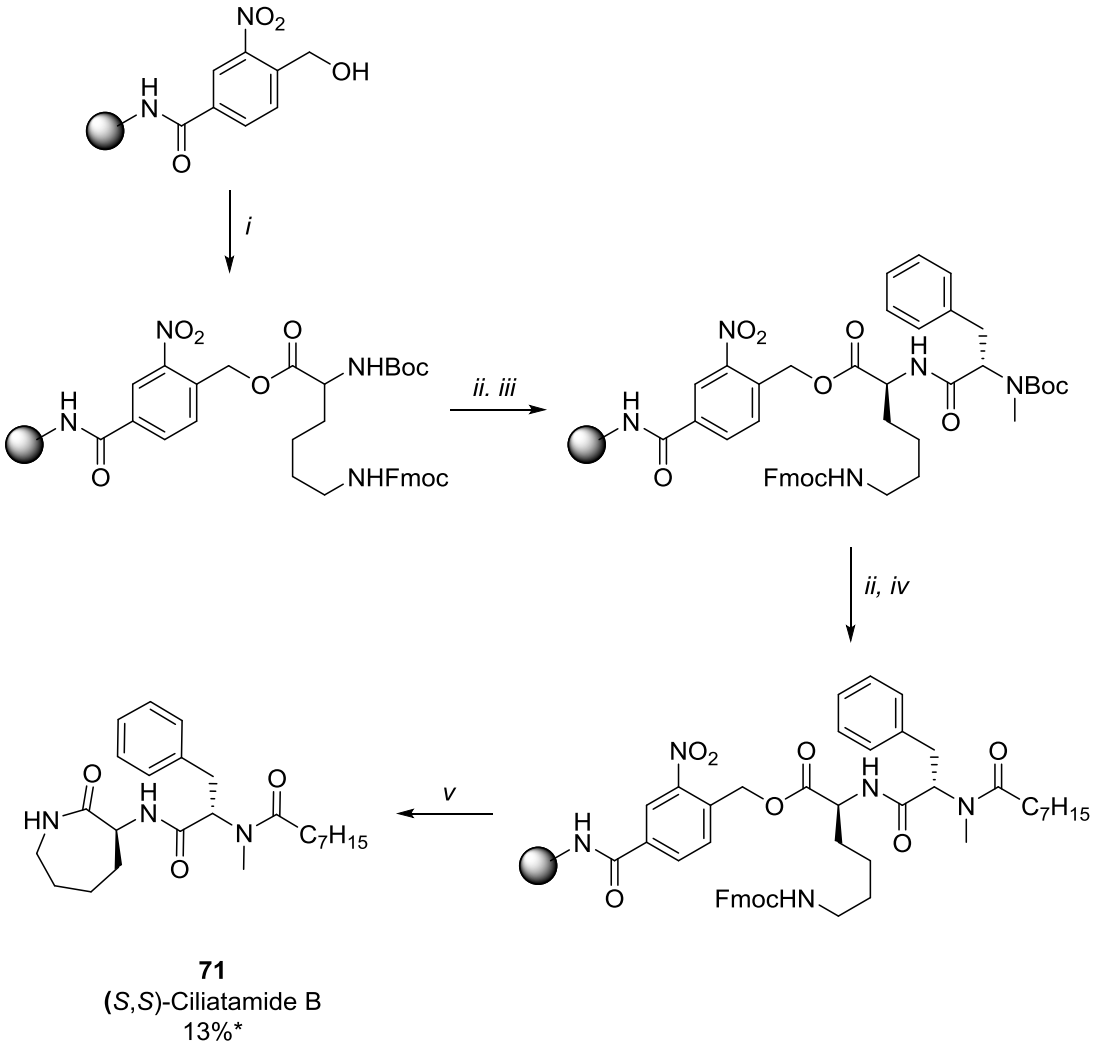


Figure A2.1: The structure of *S*,*S*-ciliatamides A-C, **70**-**72**.

A2.2 The Synthesis of Ciliatamide B

The synthesis of ciliatamide B (**71**) using "on-resin" cyclisation strategy was attempted as a trial synthesis from commercially available Boc-Lys(Fmoc)-OH, Boc-*N*-MePhe-OH and octanoic acid.

71 was synthesised in a 13% overall isolated yield after 6 on-resin steps (**Scheme A2.1**) followed by purification *via* column chromatography. This clearly shows that the scope of this strategy extends to natural products of this type and could be utilised to great effect to synthesise a huge variety of ciliatamide derivatives with relative ease.



Scheme A2.1: *i*: Boc-Lys(Fmoc)-OH, DIC, DMAP, DCM, 30 min, RT, x 2. *ii*: TFA:DCM (40% v/v), 10 min, RT, x 3. *iii*: Boc-*N*-MePhe-OH, PyBOP, DIPEA, DCM, 1 h, RT, x 2. *iv*: octanoic acid, PyBOP, DIPEA, DCM, 1 h, RT, x 2.*v*: piperidine:DMF (20% v/v) 5 min, RT; 10 min, RT.*Denotes overall yield over 6 steps.

The product purity of **70** (visualised by ^{1}H NMR, **Figure A2.2**) is tentatively deemed to be of a similar quality to the material produced by Lewis *et al.*[6] *via* solution phase synthesis. The synthesis time however is much shorter and requires only one purification step. Therefore, on-resin route is the more viable to quickly make analogues for antilieshmanial testing .

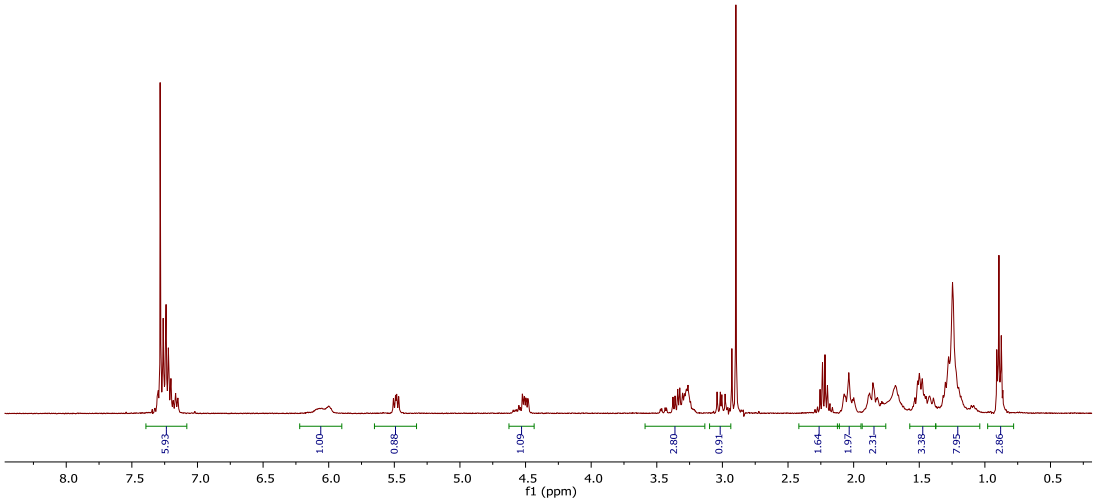


Figure A2.2: ^{1}H NMR sprectrum of *S*,*S*-ciliatamide B, **71**.

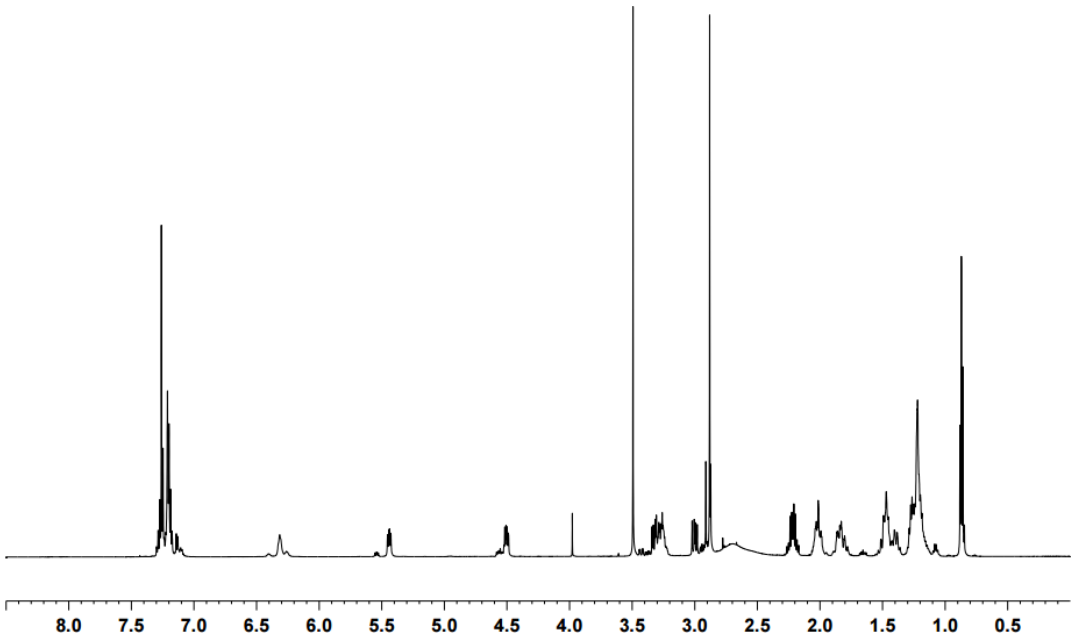


Figure A2.3: ^{1}H NMR sprectrum of *S*,*S*-ciliatamide B (**71**), reported by Lewis *et. al.*[6]

This original (*S*,*S*) chiral assignment of the natural product ciliamtamide lipopeptides by Nakao *et al.*[1] was challenged by optical rotation data obtained from chemically synthesised lipopeptides (**Figure A2.3**) by Lewis *et. al.*[6]

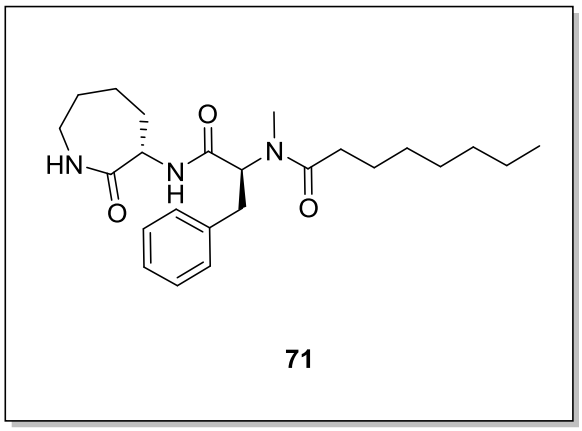
Ciliatamide B (**70**) origin	$[\alpha]^D$	Temp.	Conc.	Solvent	Ref
Isolated from: *Aaptos ciliate*	+55°	20	0.1	MeOH	1
Lewis *et al.* synthesised (*S*,*S*) diastereoisomer	-40°	20	0.1	MeOH	6
Lewis *et al.* synthesised (*R*,*R*) diastereoisomer	+49°	20	0.1	MeOH	6
On-resin synthesised (*S*,*S*) diastereoisomer	-53°	22	0.5	MeOH	

Table A2.1: Optical rotation measurements of the natural product (*S*,*S*) and (*R*,*R*) ciliamtamide B lipopeptides.

Our optical rotation data agrees with the reassignment[6] and the R,R diastereoisomer is likely to be the natural form of **70**. However, anti-parasitic testing is key to definitively understanding which is the active and hence, naturally occurring diastereoisomer.

A2.2 Experimental

(S,S) Ciliatamide B. **71**



71

To MHBA-linker resin (**65**) (1.00 eq, 0.08 mmol) a pre-mixed peptide coupling solution of Boc-L-Lys(Fmoc)-OH (3.00 eq, 0.23 mmol), DMAP (0.05 eq) and DIC (3.0 eq, 0.23 mmol) in DCM (2 mL) was added. The reaction mixture was shaken for 30 min. The solution was drained and coupling procedure was repeated. The resin was washed with DCM (5 x 5 mL). A deprotection

solution of 40% TFA in DCM (3 mL) was added to the resin and stirred for 10 min. The solution was drained and this deprotection procedure was repeated twice. The resin was washed with DCM (5 x 5 mL). A pre-mixed peptide coupling solution of Boc-L-MePhe-OH acid (3.00 eq, 0.23 mmol), DIPEA (5.00 eq, 0.39 mmol) and PyBOP (3.00 eq, 0.23 mmol) in DCM (3 mL) was added to the resin and stirred for 1 h. The solution was drained and coupling procedure was repeated. The resin was washed with DCM (5 x 5 mL) a deprotection solution of 40% TFA in DCM (3 mL) was added to the resin and stirred for 10 min. A pre-mixed peptide coupling solution of octanoic acid (3.00 eq, 0.23 mmol), DIPEA (5.00 eq, 0.39 mmol) and PyBOP (3.00 eq, 0.23 mmol) in DCM (3 mL) was added to the resin and stirred for 1 h. The solution was drained and coupling procedure was repeated. The resin was washed with DCM (5 x 5 mL) and DMF (3 x 5 mL). Finally, 20% Piperidine in DMF (3 mL) was added to the resin and stirred for 10 min at room temperature. The filtrate solution was drained, collected and then repeated once. The resin was then treated with DIPEA (10% in DCM) for 20 min at 60°C. The filtrate solution was drained, collected and then repeated once. The combined filtrate solutions from previous piperidine and DIPEA steps were evaporated under reduced pressure purified *via* preparative TLC (SiO_2; 100% EtOAc). (4 mg, 13%). $[\alpha]^{D}_{22}$ -53.0° (c 0.5, MeOH); δ_H (400 MHz; $CDCl_3$) 0.87 (3H, m, $CH_2C\mathbf{H}_3$), 1.01 to 1.34 (9H, m, CH), 1.34 to 1.54 (4H, m, CH), 1.83 (2H, m, CH_2), 2.01 (2H, m, CH_2), 2.19 (2H, m, CH_2), 2.87 (3H, m, NCH_3), 2.99 (1H, dd, *J* 10.0, 14.8, CH), 3.21 to 3.44 (3H, m, CH), 4.50 (1H, m, CH), 5.46 (1H, dd, *J* 6.2, 9.6, NH), 5.99 (1H, m, NH), 7.22 (5H, m, ArH); δ_C (100 MHz; $CDCl_3$) 175.1, 174.2, 169.7, 137.4, 129.0, 128.6, 126.7, 57.6, 52.4, 42.3, 34.4, 33.8, 31.8, 31.6, 29.4, 29.1, 28.1, 25.1, 22.7, 14.4; HRMS *m/z* (ES^+) 416.2913 ($[M+H]^+$. $C_{24}H_{38}N_3O_3$ requires 416.2913).

Data obtained is consistent with literature; see **Figure A.2.2** and **Figure A.2.3**.[6]

A2.3 References

1. Y. Nakao, S. Kawatsu, C. Okamoto, M. Okamoto, Y. Matsumoto, S. Matsunaga, R. W. M. van Soest and N. Fusetani, *J. Nat. Products*, 2008, **71**, 469-472.
2. A. H. Sharief, E. A. Gasim Khalil, T. G. Theander, A. Kharazmi, S. A. Omer and M. E. Ibrahim, *Exp. Parasitol.*, 2006, **114**, 247-252.
3. K. K. Pitzer, K. A. Werbovetz, J. J. Brendle and J. P. Scovill, *J. Med. Chem.*, 1998, **41**, 4885-4889.
4. F. Chappuis, S. Sundar, A. Hailu, H. Ghalib, S. Rijal, R. W. Peeling, J. Alvar and M. Boelaert, *Nat. Rev. Microbiol.*, 2007, **5**, 873-882.
5. M. Teixido, E. Zurita, M. Malakoutikhah, T. Tarrago and E. Giralt, *J. Am. Chem. Soc.*, 2007, **129**, 11802-11813.
6. J. A. Lewis, R. N. Daniels and C. W. Lindsley, *Org. Lett.*, 2008, **10**, 4545-4548.